Who consumes anthocyanins and anthocyanidins? Mining national retail data to reveal the influence of socioeconomic deprivation and seasonality on polyphenol dietary intake

John Harvey[1], Gavin Long[1], Roberto Mansilla[1], Simon Welham[2], Peter Rose[2], Michelle Thomas[2], Gregor Milligan[1,3], Elizabeth Dolan[1,3], Joanne Parkes[1,3], Kuzivakwashe Makokoro[1,3], James Goulding[1]

[1]N/LAB, Nottingham University Business School, University of Nottingham
[2]Division of Food, Nutrition & Dietetics, School of Biosciences, University of Nottingham
[3]Horizon Centre for Doctoral Training, University of Nottingham, UK
Corresponding author: john.harvey@nottingham.ac.uk

***Abstract*—Anthocyanins are a class of polyphenols that have received widespread recent attention due to their potential health benefits. However, estimating the dietary intake of anthocyanins at a population level is a challenging task, due to the difficulty of scaling dietary surveys. Further, there is limited evidence as to who regularly consumes anthocyanins, whether temporally, spatially, or culturally according to levels of socioeconomic deprivation. Leveraging a massive retail loyalty card dataset in the UK, we pair two years of real-world purchasing data for 619,524 regular shoppers and 207 million shopping baskets with anthocyanin estimates drawn from polyphenol databases. We subsequently analyse relative deprivation levels of the neighbourhoods in which shoppers reside, illustrating how anthocyanin intake varies according to affluence. Results indicate that deprivation is linked dramatically with both lower total intake of anthocyanins and lower breadth of dietary sources for them, potentially aggravating the incidence of diet-related diseases in the poorest sections of society.**

***Index Terms*—Anthocyanins, Anthocyanidins, Polyphenols, Socioeconomic deprivation, Food insecurity, Shopping Data, Digital Footprints**

I. Introduction

Anthocyanins, a class of polyphenol, are flavonoid compounds responsible for the red, blue, and purple colours of many fruits, vegetables, and flowers. Polyphenols have garnered recent attention within the health literature due to being ‘highly essential functional foods in our diet’ and with increasing evidence of potential health benefits, especially in relation to Non-communicable disease (NCD) [1]. However, analyses have generally been restricted to smaller cohort studies, due to the extensive logistical challenges associated with scaling dietary surveys to larger community levels.

Anthocyanins have emerged as an active area of research in the study of polyphenols. They are found in the vacuoles of plant cells and have a structure and colour that varies depending on the pH of the vacuole [2]. While the aglycon core structure of anthocyanins [3] is *anthocyanidin*, anthocyanins are glycosylated - meaning they have sugar molecules attached to them, rendering them more water-soluble and less stable than anthocyanidins.

A wide range of potential health benefits of anthocyanin consumption have been suggested in academic literature [4], [5], including but not limited to: reduction of cardiovascular disease risk [6]; reduction of cancer risk [7] [8]; improvement of cognitive health [9]; and prevention and management of diabetes and obesity [10]–[12]. Anthocyanins may also help to mitigate against the development and progression of many non-communicable and degenerative disorders. The mechanisms through which these touted benefits contribute to health remains an active and contested area of research, though some authors [13] note this may be through increasing antioxidant defences, diminishing free radical damage, reducing chronic inflammation, and minimising the risk of mutations. With anthocyanins becoming the object of increasing medical interest, attention has been drawn to exactly which people and communities might benefit the most from modifying their dietary intake. Yet, to date, there remains limited understanding and a lack of data at population levels as to *who* is consuming anthocyanins, and *where* and *when* they occur within the national diet.

A. What is known about anthocyanin consumption and how it varies by demographic factors?

Studying anthocyanin intake within broad populations is challenging for a number of reasons. Producers and manufacturers of food are not obligated to list anthocyanin content unlike standard nutritional outcomes; and polyphenols are not classified as essential nutrients. As the body does not need polyphenols to survive, they have not historically been a priority for nutritional research before 1995 [14].

The daily average intake of anthocyanins per capita can also vary dramatically across nations from several milligrams to hundreds of milligrams. Gonçalves et al. [13] note that intake is likely higher in countries with a Mediterranean diet, with abundant sources of reddish berries, red and blue-coloured fruits, and red wine. In Europe, ingestion ranges are estimated in men from 19.8 mg per day (the Netherlands) to 64.9 mg per day (Italy), and 18.4 mg per day (Spain) to 44.1 mg per day (Italy) in women - whereas in the USA, Australia, and Asian countries, the intake has been estimated at 12.5, 24.2, and 37 mg per day per person, respectively [13].

There are as yet, no recommended guidelines for anthocyanin consumption in the UK. Although there is no international alignment, the recommended daily consumption of these compounds has previously been evaluated, with China already recommending a daily intake of 50 mg per person [15]. Benchmarking daily recommendations is impeded due to the bioavailability of anthocyanins varying according to multiple factors, such as: diet, food intolerances, quantities in foods, food processing, cooking methods, and combinatorial or emergent effects of eating different food groups simultaneously [16] [17] [18].

B. Consumption of anthocyanins at socio-demographic levels

Given the challenges described above, most studies that examine consumption of anthocyanins often focus on small cohorts, using self-reported dietary surveys; or focus on small experimental studies to analyse whether anthocyanin-containing foods actually become bioavailable after ingestion. What tends to be missing from existing research is the more fundamental question of *who* actually buys foods that contain anthocyanins in the first place, and which communities and demographics might be suffering from diminished intake?

Such questions are extremely challenging to answer at both national and community levels using traditional dietary surveying techniques, due to both prohibitive costs and practical limitations [19]. However, a potential ameliorating approach lies in the interrogation of longitudinal shopping data through loyalty card records to support large-scale observational studies.

Previous research in the UK evidences that many people still experience food insecurity [20] [21], being either unable to procure enough food, or alternatively to secure a nutritionally adequate diet. Many of the people who experience food insecurity have restricted dietary choices and often have lower levels of fruit and vegetables in their diet [22] [23]. However, as anthocyanins occur naturally in many different fruits, vegetables, drinks, and alcoholic beverages, it does not necessarily follow that they would be exposed to lower levels of anthocyanins in their diet.

In view of these issues, in this study we make use of massive anonymised shopping records from a UK retailer (Co-op Food) to ask the following three research questions:

- **Which food groups are the most common sources of anthocyanins within the UK diet as observed in retail purchase data?** Further, what is the relative share of specific anthocyanins provided by these distinct food groups to the national dietary intake?
- **How does seasonal variation of produce affect the consumption of anthocyanins within the UK diet?** We know that many sources of anthocyanins are fresh fruit and vegetables. Some of these can be grown seasonally in the UK, but many are also imported throughout the year in the absence of local supply. It is not clear which specific anthocyanins in the national diet are affected by seasonal produce and which are resistant to fluctuation due to constant supply and demand.
- **What is the relationship between socioeconomic deprivation and the consumption of anthocyanins?** Lastly, we are interested in understanding whether anthocyanin consumption is affected by the relative levels of affluence across the country. Imported fruit, vegetables, and wine can all be more expensive than equivalent food groups with similar calorific and essential nutrient profiles. This begs the question of whether some consumers may be priced out of being able to purchase polyphenols such as anthocyanins, and whether any health benefits of anthocyanin consumption accrue mostly to people with greater disposable incomes.

Our approach in addressing these questions is primarily descriptive in nature and focuses on creating reliable estimates for anthocyanin/anthocyanidin content within Coop Food products, which can then be associated with other spatial, temporal and cultural variables for analysis.

II. Experimental Methodology

A. Data Sources for Anthocyanin Lookup Construction

This study examines anonymised loyalty-card shopping records from Co-op Food, one of the major grocery retailers in the UK. The chain has over 2,000 stores and has the greatest geographical coverage of any UK food retailer, with over 95 per cent of the population living within 5 miles of a store [24].

Data utilised in the study is rigorously anonymised, a process undertaken by the retailer prior to analysis, and handled in full accordance with the UK's strict confidentiality and privacy requirements[1]. With a national scope, the data set reflects grocery purchases of over 4 million consumers between 2019 and 2021. For each purchase, the data details 'when', 'where', 'what', and 'how much' each individual bought. Similarly, for each item purchased, it was possible to identify descriptors, such as the 'quantity', 'price', 'item description', 'size', and 'product category'.

Purchases of food items do not correlate exactly with the consumption of such items, and this disparity introduces noise into any analysis related to shopping data. To ensure that shopping records provide insights into the foods people actually consume, that are as representative as possible, this study focuses specifically on frequent customers. Here strict

[1] With analysis taking place under the lawful basis of GDPR 6(1)(e), 'the performance of a task in the public interest' and GDPR 9(2)(j) 'processing necessary for scientific research purposes', and all data strictly aggregated to privacy-preserving levels of total store sales.

inclusion criteria are used to select frequent customers from the initial data set. The inclusion criteria applied to each individual customer record are as follows: 1. The customer must have had at least one weekly shopping visit for at least six consecutive months; 2. The customer must have had an average spend per basket greater or equal to £5; and 3. The customer must have had an average spend per week of between £5 and £400.

A total of 619,524 frequent customers meet these criteria, reflecting purchases of more than 1,106 million items sold across 207 million transactions (i.e. number of distinct purchases or 'baskets'). The shopping data covers a time frame of 909 days from 2019-07-07 until 2021-12-31. The share of spend committed by these customers to the purchase of products containing anthocyanins is 10.4%, with over 597,365 customers (96.42% of the sample) purchasing at least one item containing anthocyanins during the time frame analysed (suggesting that most shoppers consume products with anthocyanins even if they don't necessarily do so regularly or knowingly).

Purchased products must be linked to their corresponding anthocyanin content to analyse consumption. For this anthocyanin lookup process, the open polyphenol database *Phenol-Explorer 3.0* [25] is utilised, reflecting anthocyanin content values retrieved from scientific papers published in peer-reviewed journals. Information on the original food description, the commercial origin of the sample and the number of analysed samples to produce each original content value is registered in the database, together with the analytical method used. Data are screened and, where considered acceptable, used to produce mean content values for the various anthocyanins in a range of foods and beverages.

71 distinct categories of anthocyanins and anthocyanidins identified through the Phenol Explorer database are used as the basis for lookup on products sold throughout the retailer's network. The method used to conduct the anthocyanin-food matching process is a form of 'fuzzy' text matching using trigrams in PostgreSQL. This process is performed iteratively to maximise the number of products included, regardless of product size, weight, brand, or provenance. Food products still unmatched after this step are matched manually, until the Phenol-Explorer database categories are exhausted. From the initial reference group of 71 categories, 497 food and beverage products survive, each containing at least one anthocyanin (from 59 matching categories of anthocyanins and anthocyanidins), though many contain multiple different types. Using the respective weights and sales totals of each product, sales totals are estimated for each of the 59 categories, and these statistics are linked to stores and the respective aggregate geographies in which their customers reside.

B. Feature Engineering Dependent and Independent Variables

For each research question, we are interested in assessing the amount of anthocyanins/anthocyanidins purchased by regular shoppers. In RQ1 we note a many-to-many relationship between food groups and anthocyanin compounds, motivating a focus on calculating total anthocyanin intake (mg) across core anthocyanin/anthocyanidin groups (Cyanidin, Delphinidin, Malvidin, Pelargonidin, Peonidin, Petunidin, and Pyranoanthocyanin). Each of these groups contains a range of related compounds and these are aggregated to the highest level before then studying which food groups contribute respectively to each of the anthocyanin categories.

For RQ2 we temporally aggregate each anthocyanin/anthocyanidin group according to purchases made across the UK in each month. We chose to focus on January-December 2020 and January-December 2021 to help illustrate seasonal variation across two years.

Finally, for RQ3 we aggregate the purchase histories of anthocyanins and anthocyanidins for retail customers, by the relative levels of deprivation known to exist in the neighbourhoods respective customers reside in. Each store is assigned a geospatial reference at 'Lower Super Output Area' (LSOA) level, a geographic hierarchy designed to improve the reporting of small area statistics in England and Wales (LSOAs featuring between 400 and 1,200 households). We use this geospatial reference in combination with recent data from the English Indices of Multiple Deprivation (IMD) [26] to ascertain the relative level of deprivation. IMD is a publicly available, composite measure of deprivation calculated using data on income, employment, health, education, housing, living environment and crime rate, used to rank neighbourhoods relative to one another. Using these variables, we then study the relationship between IMD-ranked deciles for customers across England and the associated average levels of anthocyanin purchased.

III. Results

A. Foods that contribute to dietary intake of anthocyanins

From the 497 matched products we found that 9 high-level food groups (as defined by Phenol Explorer) are present in the national diet of customers. We used these high-level categories as labels unless only one type of food was matched in the larger group, in which case we listed the individual food type as the label. The food groups matched across the Phenol-Explorer and the grocery retailer's products included: *Fruits - Berries*; *Fruits - Drupes*; *Fruit Juices*; *Jams - Berries*; *Black Olives*; *Leaf Vegetables*; *Pulses & Beans*; *Red Onions*; and *Wines*. A more detailed breakdown of the food groups subcategories that were matched in the database can be seen in Table 1 in the appendix, with products shown alongside the specific anthocyanin derivatives aggregated for analysis.

The food groups with the largest respective contributions to total anthocyanin/anthocyanidin intake across the sample were *Wines* (Malvidin = 36.35%; Petunidin = 5.45%; Peonidin = 4.18%; and Pyranoanthocyanin 1.4%) and *Berries* (Pelargonidin, 19.82%; Delphinidin = 4.96%; and Cyanidin = 4.3%). Figure 1 illustrates the many-to-many relationship detected between 7 categories of anthoncyanin/anthocyanidin intake and the relative share that the 9 high-level food groups contribute across the population.

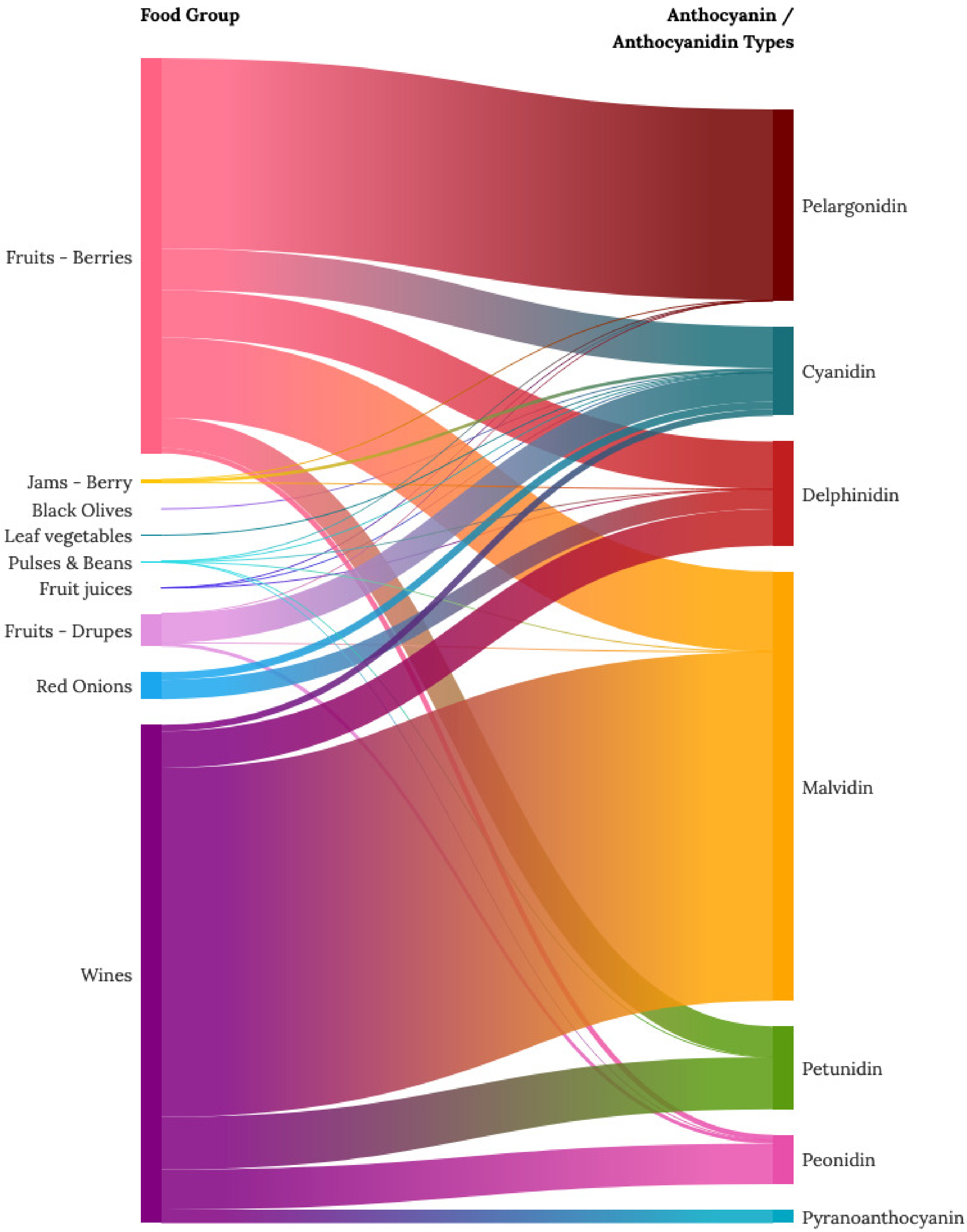

Fig. 1: Sankey diagram showing food categories purchased by 619,524 regular customers and their associated contributions to distinct groups of anthocyanins and anthocyanidins. Food categories are shown by size according to total weight purchased (grams) and anthocyanin/anthocyanidin categories are shown according to mg derived from each respective food source

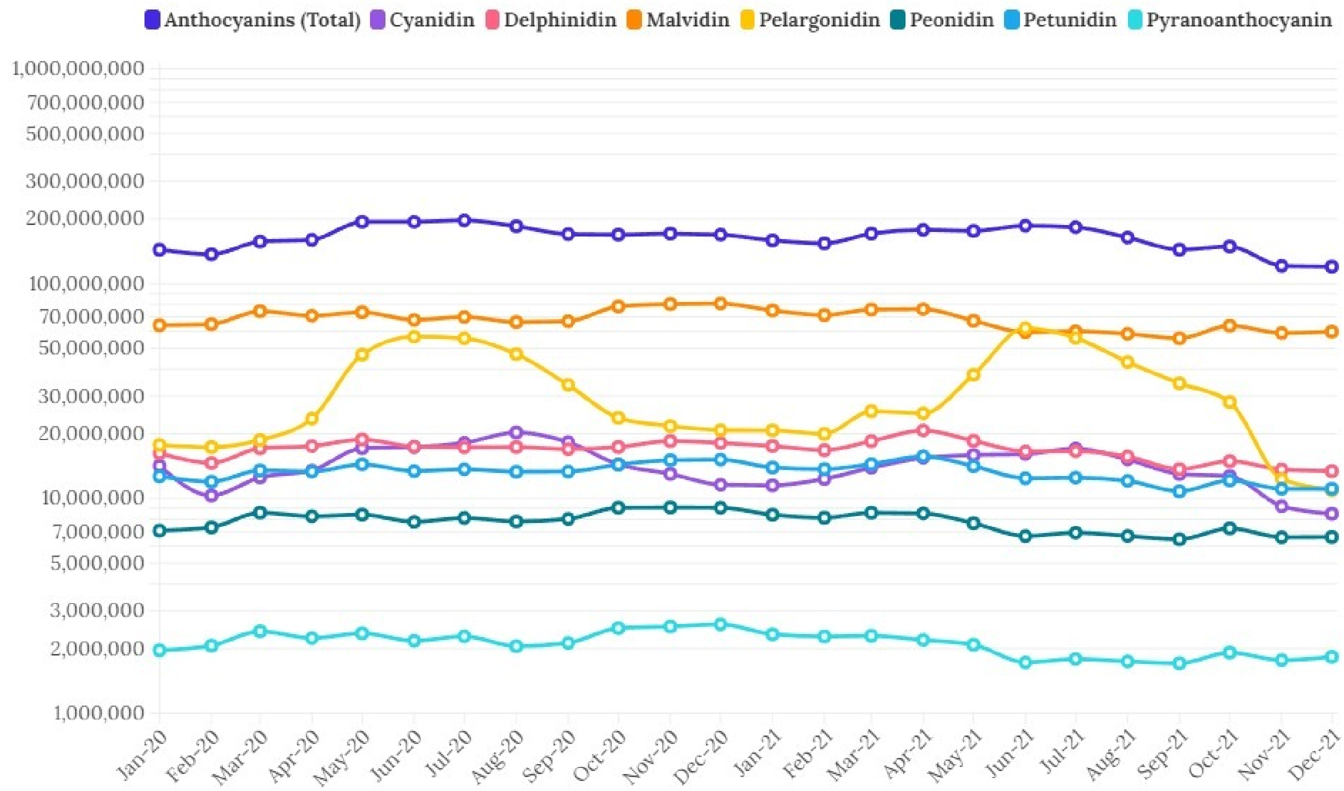

Fig. 2: Seasonal variation in the purchasing habits of 619,524 regular shoppers buying anthocyanin/anthocyanidin (mg) containing products from January 2020 to December 2021. Results are shown on log scale

B. Seasonal variation in the purchasing habits of people buying anthocyanin/anthocyanidin containing products

Figure 2 depicts the seasonal variation of anthocyanin and anthocyanidin consumption from January 2020 until December 2021 for the entire sample. Given that wine contributes 51% to the total dietary anthocyanin intake of the sample, we see that for many categories there is a fairly consistent and stable purchasing rate across time (including during peak purchasing during the 2020 COVID-19 lockdowns). There is, therefore, a segment of anthocyanin types which are relatively season-resistant, particularly those that are prevalent in *red* wine (i.e. *Malvidin*, *Petunidin*, *Peonidin*, and especially *Pyranoanthocyanins* which are formed by yeast during fermentation processes or during controlled oxygenation techniques during the red wine ageing process [27] [28]).

By contrast we see far greater seasonal variation for *Pelargonidin*, which is overwhelmingly derived from berries, many of which grow seasonally in the UK, such as strawberry, raspberry, and blackcurrant. Similarly, *Cyanidin* experiences seasonal variation as it derives principally from berries and drupes, some of which are grown widely in the UK during warmer months, such as plums and cherries.

C. Socioeconomic deprivation and dietary intake of anthocyanins/anthocyanidins

A linear relationship is clearly evident between the average consumption of anthocyanins and the deprivation deciles (according to the Index of Multiple Deprivation) of the area in which shoppers reside, as is shown clearly in Figure 3. For each category of anthocyanidin/anthocyanin we find that the R^2 demonstrates the influence of deprivation on dietary intake: Total anthocyanins ($R^2 = 0.94$), Cyanidin ($R^2 = 0.95$), Delphinidin ($R^2 = 0.95$), Malvidin ($R^2 = 0.92$), Pelargonidin ($R^2 = 0.97$), Peonidin ($R^2 = 0.91$), Petunidin ($R^2 = 0.94$), and Pyranoanthocyanin ($R^2 = 0.91$).

To help illustrate the potential impact of this linear correlation, we can compare the intakes of shoppers who reside in the most deprived areas (IMD 1) and the shoppers who reside in the least deprived areas (IMD 10). Across all measures of anthocyanin intake we see notable increases when comparing IMD1 and IMD10: Total anthocyanins (+118.37%), Cyanidin (+97.44%), Delphinidin (+134.15%), Malvidin (+132.29%), Pelargonidin (+86.90%), Peonidin (+120.60%), Petunidin (+140.60%), and Pyranoanthocyanin (+120.60%). These results provide first evidence at a national scale indicating that the experience of deprivation (at least when modelled as a composite variable) is strongly related to the dietary intake of anthocyanins. This result is perhaps most clearly illustrated

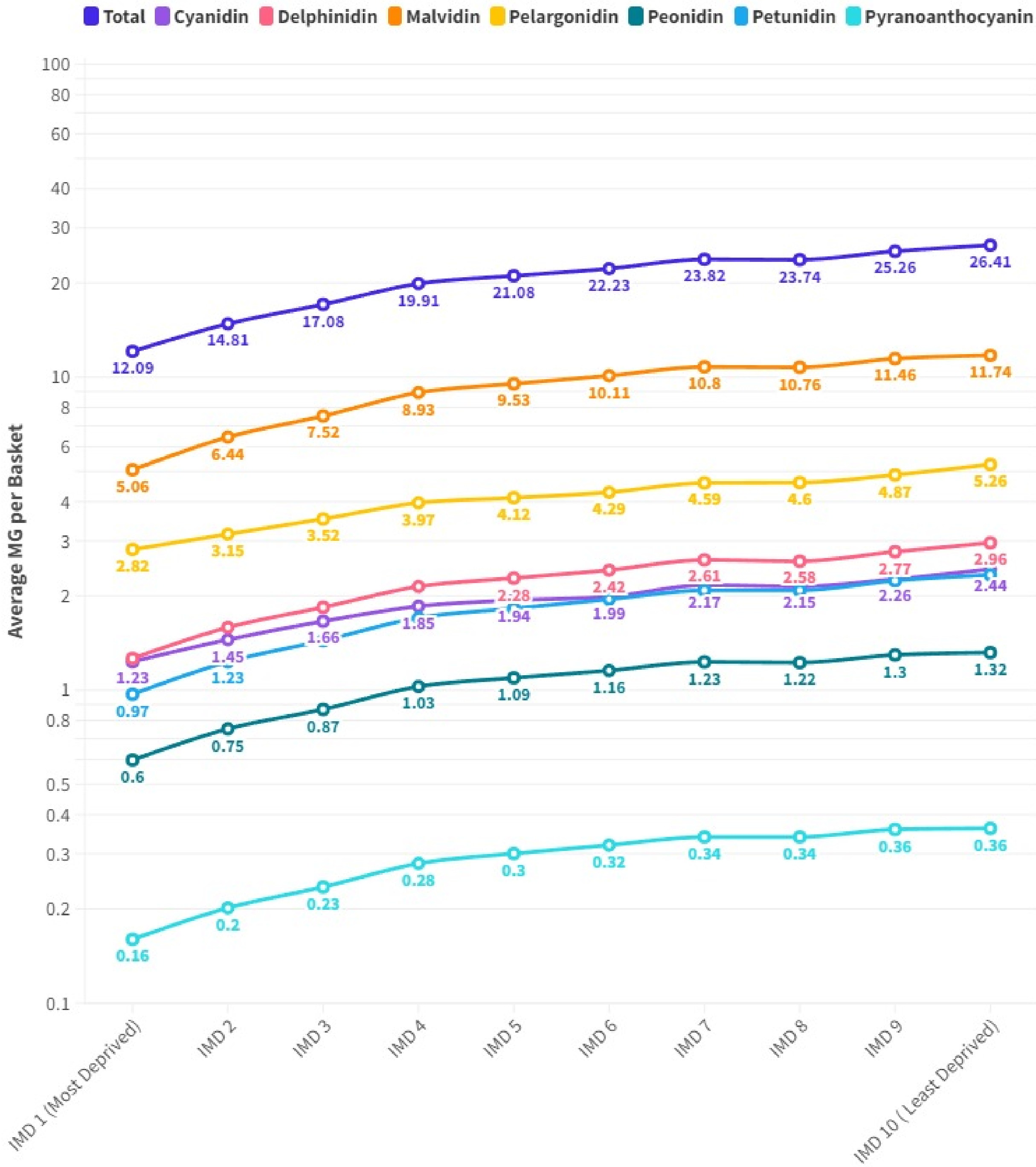

Fig. 3: Average anthocyanin MG per basket, grouped by relative levels of deprivation according to the Index of Multiple Deprivation. Results are shown on log scale.

when considering *Petunidin*, which shows the strongest effect and is almost solely derived from red wine and blueberries, both of which are products that are relatively expensive to UK consumers due to the limited scale of national production, therefore, the essential reliance on importing.

IV. Discussion

To the best of our knowledge, this initial exploratory study reflects the first attempt to pair national level retail 'Big data' with polyphenol databases to provide estimates of dietary intake. Though the approach provides an unprecedented scale of estimation and a strong signal relating deprivation deciles to polyphenol intake, it nonetheless has several notable limitations. First, there are likely many processed foods with anthocyanin-containing ingredients that at present, are impossible to evaluate using this method. This is primarily a shortcoming of what is currently known about how food processing affects the prevalence of anthocyanins, and indeed, polyphenols more generally.

Environmental and genetic factors influencing the uptake of anthocyanin, post-harvest conditions and preparation methods drive differences in the levels of these molecules in plant tissues. It is widely known that temperature drives loss in anthocyanin levels in food products [29]. Heat processing, a method widely used in food preservation can cause changes in colour pigmentation [30] [31]. Currently, cool temperatures favour anthocyanin stability in fruits and anthocyanin-containing products [32]. Aside from temperature and oxygen levels [33], light [34], food preservatives like sulphites [35], and the presence of other antioxidants such as ascorbic acid [36], can all affect the stability of plant derived anthocyanin associated compounds in food products (reviewed in 'Anthocyanins: Factors affecting their stability and degradation' [37]).

Taken together, it is clear that several factors influence the levels of anthocyanins present in the diets of humans. Moving forward, it will be important to consider these factors in studies focused on anthocyanins in the human food chain. Similarly, open databases have captured limited information on how anthocyanins may vary, as a consequence of the seasonality and local growing conditions of where fruits and vegetables are produced. Naturally, this may in some unknown instances lead to systematic under/over-estimation. Further research to extend open-access datasets would help to provide greater accuracy of intake estimates. Additionally, it would be useful to augment retail data with information on food consumed outside of the home, for example restaurants, cafes, and in the case of berries, estimates on foraging, as each of these inevitably contribute somewhat to overall intake.

Though we do not attempt to provide daily intake recommendations as described in the introduction, instead providing analysis at the basket level, it is clear that many daily averages provided in previous works are inevitably concealing heterogeneous populations that vary massively in their intakes. Historically, the conceptualisation of food and nutrition security has focused on whether people can acquire enough calories and whether the food they obtain contains enough essential nutrients. However, if consuming polyphenols such as anthocyanins provides beneficial health effects and reduction in diet-related diseases, it may well be worth revisiting insecurity categorisation, particularly where the protective effects of food are divided along traditional socioeconomic fault lines of income, class, and privilege. We have shown for anthocyanin/anthocyanidin intakes that there is a clear disparity between the rich and the poor, but future research should address whether this same disparity holds across multiple phenol categories, and whether a broader category of *'polyphenol insecurity'* exists nationally and internationally. Furthermore, in future research, it will be crucial to consider adjusting analyses to account for confounding factors associated with poor health and poverty. These factors include smoking, alcohol consumption, and inadequate diet. Interestingly, these particular confounding variables have the potential to be measured and assessed via the same loyalty card data which could also be used to account for the potential protective effects of other food categories alongside polyphenols, for example, dairy and fibre-rich diets [38], [39].

Moreover, this study raises questions about the provenance of anthocyanins in the national diet and whether year-round access to these compounds through imported products, such as berries and wine, may actually exacerbate existing international climate inequalities, while simultaneously only providing health benefits for a select group of people able to afford them.

The method pioneered in this study provides a preliminary insight into what is potentially possible for the scientific study of polyphenols through digital footprint data such as loyalty card records. We believe the most 'fruitful' application of this approach would be to pair large cohort studies focused on measures of health and disease with longitudinal shopping data. For those health conditions purported to benefit from anthocyanin consumption, especially those that are diet-related diseases, there are clear opportunities to yield otherwise impossible insights.

V. Acknowledgements

This work was partly supported by the Future Food Beacon at University of Nottingham, the Horizon Centre for Doctoral Training at the University of Nottingham (UKRI grant EP/S023305/1) and EPSRC grants (EP/V053922/1 and EP/V51973X/1): 'the CIVIC Project: A Sustainable Platform for syndromic-surveillance via Health, Deprivation and Mass Loyalty-Card Datasets'.

References

[1] Ananya Rana, Mrinal Samtiya, Tejpal Dhewa, Vijendra Mishra, and Rotimi E Aluko. Health benefits of polyphenols: A concise review. *Journal of Food Biochemistry*, 46(10):e14264, 2022.

[2] Elisa Petrussa, Enrico Braidot, Marco Zancani, Carlo Peresson, Alberto Bertolini, Sonia Patui, and Angelo Vianello. Plant flavonoids—biosynthesis, transport and involvement in stress responses. *International journal of molecular sciences*, 14(7):14950–14973, 2013.

[3] Francesca Giampieri, Danila Cianciosi, José M Alvarez-Suarez, José L Quiles, Tamara Y Forbes-Hernández, María D Navarro-Hortal, Michele Machì, Michele Ramón del Jesús Palí Casanova, Julio César Martínez Espinosa, Xiumin Chen, et al. Anthocyanins: what do we know until now? *Journal of Berry Research*, (Preprint):1–6, 2023.

[4] Hock Eng Khoo, Azrina Azlan, Sou Teng Tang, and See Meng Lim. Anthocyanidins and anthocyanins: Colored pigments as food, pharmaceutical ingredients, and the potential health benefits. *Food & nutrition research*, 61(1):1361779, 2017.

[5] Elisa Pojer, Fulvio Mattivi, Dan Johnson, and Creina S Stockley. The case for anthocyanin consumption to promote human health: a review. *Comprehensive reviews in food science and food safety*, 12(5):483–508, 2013.

[6] Sonia de Pascual-Teresa, Diego A Moreno, and Cristina García-Viguera. Flavanols and anthocyanins in cardiovascular health: a review of current evidence. *International journal of molecular sciences*, 11(4):1679–1703, 2010.

[7] Li-Shu Wang and Gary D Stoner. Anthocyanins and their role in cancer prevention. *Cancer letters*, 269(2):281–290, 2008.

[8] Bo-Wen Lin, Cheng-Chen Gong, Hai-Fei Song, and Ying-Yu Cui. Effects of anthocyanins on the prevention and treatment of cancer. *British journal of pharmacology*, 174(11):1226–1243, 2017.

[9] Nathalia Medina dos Santos, Patricia Berilli Batista, Angela Giovana Batista, and Mário Roberto Maróstica Júnior. Current evidence on cognitive improvement and neuroprotection promoted by anthocyanins. *Current opinion in food science*, 26:71–78, 2019.

[10] Yoon-Mi Lee, Young Yoon, Haelim Yoon, Hyun-Min Park, Sooji Song, and Kyung-Jin Yeum. Dietary anthocyanins against obesity and inflammation. *Nutrients*, 9(10):1089, 2017.

[11] Elena Azzini, Jasminka Giacometti, Gian Luigi Russo, et al. Antiobesity effects of anthocyanins in preclinical and clinical studies. *Oxidative medicine and cellular longevity*, 2017, 2017.

[12] Honghui Guo and Wenhua Ling. The update of anthocyanins on obesity and type 2 diabetes: experimental evidence and clinical perspectives. *Reviews in Endocrine and Metabolic Disorders*, 16:1–13, 2015.

[13] Ana C Gonçalves, Ana R Nunes, Amílcar Falcão, Gilberto Alves, and Luís R Silva. Dietary effects of anthocyanins in human health: A comprehensive review. *Pharmaceuticals*, 14(7):690, 2021.

[14] Augustin Scalbert, Ian T Johnson, and Mike Saltmarsh. Polyphenols: antioxidants and beyond. *The American journal of clinical nutrition*, 81(1):215S–217S, 2005.

[15] Chinese Nutrition Society. Chinese dris handbook. *Standards Press of China*, 2013.

[16] Jim Fang. Bioavailability of anthocyanins. *Drug metabolism reviews*, 46(4):508–520, 2014.

[17] Anna Rafaela Cavalcante Braga, Daniella Carisa Murador, Leonardo Mendes de Souza Mesquita, and Veridiana Vera de Rosso. Bioavailability of anthocyanins: Gaps in knowledge, challenges and future research. *Journal of Food Composition and Analysis*, 68:31–40, 2018.

[18] Iva Fernandes, Ana Faria, Conceição Calhau, Victor de Freitas, and Nuno Mateus. Bioavailability of anthocyanins and derivatives. *Journal of functional foods*, 7:54–66, 2014.

[19] Victoria L Jenneson, Francesca Pontin, Darren C Greenwood, Graham P Clarke, and Michelle A Morris. A systematic review of supermarket automated electronic sales data for population dietary surveillance. *Nutrition reviews*, 80(6):1711–1722, 2022.

[20] Georgiana Nica-Avram, John Harvey, Gavin Smith, Andrew Smith, and James Goulding. Identifying food insecurity in food sharing networks via machine learning. *Journal of Business Research*, 131:469–484, 2021.

[21] Georgiana Nica-Avram, John Harvey, James Goulding, Benjamin Lucas, Andrew Smith, Gavin Smith, and Bertrand Perrat. Fims: Identifying, predicting and visualising food insecurity. In *Companion Proceedings of the Web Conference 2020*, pages 190–193, 2020.

[22] Narges Ebadi and Davod Ahmadi. Association between food insecurity and low fruit and vegetable intake in the uk in 2016: a cross-sectional study. *The Lancet*, 394:S36, 2019.

[23] O Turnbull, M Homer, and H Ensaff. Food insecurity: Its prevalence and relationship to fruit and vegetable consumption. *Journal of human nutrition and dietetics*, 34(5):849–857, 2021.

[24] Tim Rains and Paul Longley. The provenance of loyalty card data for urban and retail analytics. *Journal of Retailing and Consumer Services*, 63:102650, 2021.

[25] Joseph A Rothwell, Jara Perez-Jimenez, Vanessa Neveu, Alexander Medina-Remon, Nouha M'hiri, Paula García-Lobato, Claudine Manach, Craig Knox, Roman Eisner, David S Wishart, et al. Phenol-explorer 3.0: a major update of the phenol-explorer database to incorporate data on the effects of food processing on polyphenol content. *Database*, 2013:bat070, 2013.

[26] Ministry of Housing, Communities & Local Government. English indices of deprivation 2019.

[27] Jingren He, Celestino Santos-Buelga, Nuno Mateus, and Victor de Freitas. Isolation and quantification of oligomeric pyranoanthocyanin-flavanol pigments from red wines by combination of column chromatographic techniques. *Journal of Chromatography A*, 1134(1-2):215–225, 2006.

[28] Vessela Atanasova, Hélène Fulcrand, Véronique Cheynier, and Michel Moutounet. Effect of oxygenation on polyphenol changes occurring in the course of wine-making. *Analytica Chimica Acta*, 458(1):15–27, 2002.

[29] Iuliana Aprodu, Ștefania Adelina Milea, Elena Enachi, Gabriela Râpeanu, Gabriela Elena Bahrim, and Nicoleta Stănciuc. Thermal degradation kinetics of anthocyanins extracted from purple maize flour extract and the effect of heating on selected biological functionality. *Foods*, 9(11):1593, 2020.

[30] Delia B Rodriguez-Amaya. Update on natural food pigments-a mini-review on carotenoids, anthocyanins, and betalains. *Food Research International*, 124:200–205, 2019.

[31] GH Laleh, H Frydoonfar, R Heidary, R Jameei, and S Zare. The effect of light, temperature, ph and species on stability of anthocyanin pigments in four berberis species. *Pakistan Journal of Nutrition*, 5(1):90–92, 2006.

[32] Jarkko Hellström, Pirjo Mattila, and Reijo Karjalainen. Stability of anthocyanins in berry juices stored at different temperatures. *Journal of Food Composition and Analysis*, 31(1):12–19, 2013.

[33] Negese Yazie Amogne, Delele Worku Ayele, and Yeshitila Asteraye Tsigie. Recent advances in anthocyanin dyes extracted from plants for dye sensitized solar cell. *Materials for Renewable and Sustainable Energy*, 9:1–16, 2020.

[34] Remini Hocine, Dahmoune Farid, Sahraoui Yasmine, Madani Khodir, VN Kapranov, and EF Kiselev. Recent advances on stability of anthocyanins. *. :* , 13(4):257–286, 2018.

[35] Rodrigo N Cavalcanti, Diego T Santos, and Maria Angela A Meireles. Non-thermal stabilization mechanisms of anthocyanins in model and food systems—an overview. *Food research international*, 44(2):499–509, 2011.

[36] Jacob E Farr and M Monica Giusti. Investigating the interaction of ascorbic acid with anthocyanins and pyranoanthocyanins. *Molecules*, 23(4):744, 2018.

[37] Bianca Enaru, Georgiana Drețcanu, Teodora Daria Pop, Andreea Stănilă, and Zorița Diaconeasa. Anthocyanins: Factors affecting their stability and degradation. *Antioxidants*, 10(12):1967, 2021.

[38] Laura Barrubés, Nancy Babio, Nerea Becerra-Tomás, Núria Rosique-Esteban, and Jordi Salas-Salvadó. Association between dairy product consumption and colorectal cancer risk in adults: a systematic review and meta-analysis of epidemiologic studies. *Advances in Nutrition*, 10(suppl_2):S190–S211, 2019.

[39] Dagfinn Aune, Doris SM Chan, Rosa Lau, Rui Vieira, Darren C Greenwood, Ellen Kampman, and Teresa Norat. Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies. *Bmj*, 343, 2011.

APPENDIX

Please see overleaf.

Product Category	**Product Subcategory**
Fruits-Berries	Blackberry European Cranberry Black Grape Green Grape Highbush Blueberry Red Raspberry Strawberry
Black Olives	Black Olive
Fruit juices	Pomegranate
Fruits-Drupes	Nectarine Peach Plum Sweet Cherry
Jams-Berry	Blackcurrant Red Raspberry Strawberry
Leaf vegetables	Red Lettuce
Red Onions	Red Onion
Pulses-Beans	Black Common Bean
Wines	Red Wine; White Wine

Anthocyanin / Anthocyanidins Category	**Anthocyanin / Anthocyanidins Subcategory**
Cyanidin	Cyanidin; Cyanidin 3-O-arabinoside; Cyanidin 3-O-xyloside; Cyanidin 3-O-galactoside; Cyanidin 3-O-(6”-malonyl-3”-glucosyl-glucoside); Cyanidin 3-O-(6”-dioxalyl-glucoside); Cyanidin 3-O-rutinoside; Cyanidin 3-O-(6”-succinyl-glucoside); Cyanidin 3-O-(6”-malonyl-glucoside); Cyanidin 3-O-(6”-acetyl-glucoside); Cyanidin 3-O-(6”-acetyl-galactoside); Cyanidin 3,5-O-diglucoside; Cyanidin 3-O-(6”-p-coumaroyl-glucoside); Cyanidin 3-O-sophoroside; Cyanidin 3-O-glucosyl-rutinoside; Cyanidin 3-O-glucoside
Delphinidin	Delphinidin 3-O-glucoside; Delphinidin 3-O-feruloyl-glucoside; Delphinidin 3-O-(6”-acetyl-galactoside); Delphinidin 3-O-arabinoside; Delphinidin 3-O-(6”-p-coumaroyl-glucoside); Delphinidin 3-O-rutinoside; Delphinidin 3-O-(6”-acetyl-glucoside); Delphinidin 3-O-glucosyl-glucoside; Delphinidin 3-O-galactoside; Delphinidin 3,5-O-diglucoside
Pelargonidin	Pelargonidin; Pelargonidin 3,5-O-diglucoside; Pelargonidin 3-O-(6”-malonyl-glucoside); Pelargonidin 3-O-(6”-succinyl-glucoside); Pelargonidin 3-O-glucoside; Pelargonidin 3-O-arabinoside; Pelargonidin 3-O-rutinoside
Malvidin	Malvidin 3-O-(6”-p-coumaroyl-glucoside); Malvidin 3-O-glucoside; Malvidin 3-O-(6”-acetyl-galactoside); Malvidin 3-O-galactoside; Malvidin 3-O-(6”-caffeoyl-glucoside); Malvidin 3-O-(6”-acetyl-glucoside); Malvidin 3-O-arabinoside; Malvidin 3,5-O-diglucoside
Petunidin	Petunidin 3-O-(6”-acetyl-galactoside Petunidin 3-O-(6”-p-coumaroyl-glucoside) Petunidin 3-O-galactoside; Petunidin 3-O-glucoside Petunidin 3-O-(6”-acetyl-glucoside)
Peonidin	Peonidin; Peonidin 3-O-(6”-acetyl-glucoside); Peonidin 3-O-galactoside; Peonidin 3-O-(6”-acetyl-galactoside); Peonidin 3-O-rutinoside; Peonidin 3-O-(6”-p-coumaroyl-glucoside); Peonidin 3-O-glucoside; Peonidin 3-O-arabinoside
Pyranoanthocyanin	Pinotin A Vitisin A Pigment A

TABLE I: Category and subcategories of food, anthocyanidins and anthocyanin derivatives matched for analysis.