nature food

Article

https://doi.org/10.1038/s43016-025-01214-y

Biodiverse diets present co-benefits for greenhouse gas emissions, land use, mortality rates and nutritional adequacy in Europe

Received: 11 October 2024

Accepted: 11 July 2025

Published online: 20 August 2025

Check for updates

A list of authors and their affiliations appears at the end of the paper

Dietary diversity is vital for public health nutrition, yet the co-benefits of increasing dietary species richness (DSR) on human and environmental health remain unassessed. Here we explore associations between DSR and greenhouse gas emissions, land use, nutrient adequacy and mortality rates among European Investigation into Cancer and Nutrition (EPIC) study participants. Total DSR was positively associated with probability of adequate nutrient intake diet scores and inversely related to mortality rates; similar results were observed for plant DSR. Animal DSR was inversely associated with probability of adequate nutrient intake diet scores and neutrally associated with mortality rates. Neutral associations for total DSR and positive associations for animal DSR were found with greenhouse gas emissions and land use. Conversely, plant DSR was inversely associated with greenhouse gas emissions and land use. These findings from Europe suggest modest benefits of dietary plant biodiversity for nutrient adequacy and environmental health, with stronger inverse associations with mortality rates, while highlighting the potential adverse environmental impacts of diets rich in animal-sourced foods.

Diets are pivotal in shaping environmental and human health[1]. Agrifood systems are the primary driver of global biodiversity loss, land use and elevated dietary greenhouse gas (GHG) emissions[2,3]. Over recent decades, diets worldwide have undergone profound changes, marking a nutrition transition characterized by disparities in access to and consumption of a diversity of micronutrient-rich foods[4]. This is juxtaposed with the widespread availability of energy-dense and highly processed alternatives. These dietary shifts have resulted in a rise in the incidences of chronic non-communicable diseases[5]. This transition is skewed towards food production systems based on a narrow range of crops and animal species, leading to homogenized and non-diverse diets, as well as increased biodiversity loss[6–8]. Simultaneously, food production exerts substantial pressure on lands and freshwater resources, pushing the planet beyond its boundaries[9]. Proactive measures are imperative to safeguard the well-being of future generations, calling for a shift in lifestyle behaviours and a move towards diets that are both nutritious and environmentally sustainable. Such diets must work synergistically, by mitigating dietary GHG emissions and land use while reducing malnutrition in all its forms and improving human health and ecosystem resilience[10,11].

Food biodiversity, which encompasses the diversity of wild or cultivated plants, animals and other organisms, used for foods and drinks[12], could serve as a cross-cutting link between public health, environmental health and biodiversity stewardship, complementing existing dietary recommendations that promote sustainable healthy diets[3]. While biodiverse diets, consisting of a wide diversity of biological species, hold the potential to promote ecosystem resilience and sustainability, their impact on the environment largely depends on the types of food and where and how those foods are produced, with monoculture cropping practices and concentrated animal feeding operations posing major risks to local biodiversity[10,11]. Dietary diversity often leads to greater nutrient intakes and positive nutritional outcomes but typically refers to diversity between food groups and rarely considers diversity within food groups or intakes of distinct species[13].

e-mail: jeroen.berden@ugent.be

While many metrics have been tested to quantify food biodiversity, dietary species richness (DSR) remains a promising measure because of its straightforward interpretation, construct validity and comparability across contexts[14]. Evidence has shown DSR to be positively associated with micronutrient adequacy in low- and middle-income countries and is linked to lower mortality and reduced rates of gastrointestinal cancers in Europe[6,15,16]. However, previous studies have focused on total DSR without considering potential differences in associations for plant and animal species, thus limiting the depth of insight into these relationships. Furthermore, no studies have assessed the environmental implications of recommending greater food biodiversity, although understanding these impacts is crucial for aligning dietary recommendations with the Sustainable Development Goals and Kunming–Montreal Global Biodiversity Framework[17,18].

Although previous research has highlighted the co-benefits of sustainable healthy diets, such as the EAT–Lancet planetary health diet[19], evidence is limited for food biodiversity. Understanding the impact of greater DSR on human and environmental health is crucial for developing sustainable public health policies, as current guidelines for healthy diets may not go hand in hand with improved environmental sustainability. This study addresses these gaps by investigating the relationship between DSR and mortality, nutrient adequacy and environmental impacts. Specifically, we examine these relationships for total DSR, as well as DSR_{Plant} and DSR_{Animal} separately, using data from a large and diverse pan-European cohort.

Results

Baseline characteristics

The 417,423 participants included in this study of the European Investigation into Cancer and Nutrition (EPIC) cohort had a mean (s.d.) age at recruitment of 51.4 (10) years. Furthermore, 285,641 (68%) individuals were female and most of the participants were from the UK (18%), France (16%) and Denmark (12%). More than half of the participants (58%) were married or living together, and 29% did not attain an educational degree or completed primary school only. Half of the participants (50%) had never smoked before and 55% reported being physically inactive. The mean (s.d.) number of species consumed per person per year (DSR_{Total}) was 66.5 (15.5), with DSR_{Animal} and DSR_{Plant} being 12.5 (4.3) and 47.2 (12.1), respectively. The mean (s.d.) of dietary GHG emissions was 5.2 kg carbon dioxide equivalent (CO_2e) per kg of food per day (1.82), while land use was 6.7 m^2 per year per kg of food per day (2.60). The mean (s.d.) probability of adequate nutrient intake diet (PANDiet) score, including both macronutrients and micronutrients ($PANDiet_{Overall}$), was 61.8% (8.3) and micronutrient adequacy ($PANDiet_{Micro}$) was 67.7% (15.8) (Table 1).

Associations between DSR and mortality rates

In total, 45,489 deaths were recorded from all causes with a median follow-up time of 17 years. Multivariable-adjusted analysis indicated that higher DSR_{Total} and DSR_{Plant} were both inversely associated with all-cause mortality. For DSR_{Total}, comparing the fifth with the first quintile, the adjusted hazard ratio (HR) for all-cause mortality was 0.62 (95% confidence interval (CI) 0.59 to 0.65). Overall, a dose–response association was observed across quintiles. One-s.d. increments in DSR_{Total} and DSR_{Plant} were associated with lower all-cause mortality rates (HR 0.84; 95% CI 0.82 to 0.85 and HR 0.84; 95% CI 0.82 to 0.86, respectively). For DSR_{Animal}, no associations were observed with death, except for a weak negative association when comparing the fifth with the first quintile (HR 0.94; 95% CI 0.89 to 0.98) (Fig. 1). Associations with cause-specific mortality, including cancer, cardiovascular disease, coronary heart disease, respiratory disease and digestive disease, were similar. DSR_{Plant} was positively associated with lower rates of cause-specific mortality rates, for 1-s.d. increments and across quintiles, whereas relationships for DSR_{Animal} were neutral (Supplementary Table 4).

The relationship between DSR and all-cause mortality was also quantified using the rate advancement period (RAP) expressed in years, which were all found to be negative, indicating a delayed risk of mortality (Fig. 1). For DSR_{Total}, the RAP comparing the fifth to the first quintile was −4.83 years (95% CI −6.46 to −3.20), while for DSR_{Plant} the RAP was −4.51 years (95% CI −6.04 to −2.99). For a 1-s.d. increment in DSR_{Total} and DSR_{Plant}, the RAP estimates were −1.76 years (95% CI −1.93 to −1.59) and −1.76 years (95% CI −1.97 to −1.54), respectively.

Associations between DSR and nutrient adequacy

In adjusted linear models, DSR_{Total} was positively associated with $PANDiet_{Overall}$ and $PANDiet_{Micro}$, with regression coefficients (β) for a 1-s.d. increment being 0.29 percentage points (95% CI 0.26 to 0.33) and 1.64 percentage points (95% CI 1.60 to 1.69), respectively. Positive associations were also found for a 1-s.d. increment in DSR_{Animal} and $PANDiet_{Micro}$ (0.59 percentage points; 95% CI 0.55 to 0.63) and between DSR_{Plant} and $PANDiet_{Micro}$ (1.22 percentage points; 95% CI 1.16 to 1.28). Conversely, a 1-s.d. increment in DSR_{Animal} showed inverse associations with $PANDiet_{Overall}$ (−0.09 percentage points; 95% CI −0.12 to −0.06) (Fig. 2).

Nutrient adequacy as mediator of DSR and all-cause mortality

For DSR_{Total}, $PANDiet_{Overall}$ mediated only 1.08% (95% CI 0.54 to 1.42) of the association with lower all-cause mortality rates, while $PANDiet_{Micro}$ mediated 4.69% (95% CI 3.05 to 5.73). Similarly, for DSR_{Plant}, $PANDiet_{Overall}$ and $PANDiet_{Micro}$ mediated only 1.53% (95% CI 0.77 to 2.20) and 3.63% (95% CI 2.6 to 4.74) of the inverse association with death. For DSR_{Animal}, no significance with mortality was observed (Supplementary Table 5).

Associations between DSR and environmental impacts

In adjusted linear models, DSR_{Total} did not show clear patterns of association with environmental impacts. However, DSR_{Animal} was positively associated with GHG emissions (1-s.d. increment: 0.06 $kgCO_2e$ per day; 95% CI 0.05 to 0.06) and land use (0.07 m^2 per day; 95% CI 0.07 to 0.08). By contrast, negative associations were found between DSR_{Plant} and GHG emissions and land use in the total sample (Fig. 2).

The machine learning models, using Shapley additive explanations (SHAP) values of PANDiet scores, GHG emissions and land use, concurred with the observed relationships derived from the linear models (Fig. 3). Positive rank correlations were found for DSR_{Total} and $PANDiet_{Overall}$ ($\rho = 0.26$) and DSR_{Plant} and $PANDiet_{Overall}$ ($\rho = 0.62$), while inverse correlations were observed for DSR_{Animal} and $PANDiet_{Overall}$ ($\rho = -0.79$). For land use and GHG emissions, a positive correlation was reported with DSR_{Animal} ($\rho = 0.63$ and $\rho = 0.56$, respectively). Conversely, negative correlations were observed between DSR_{Plant} and land use ($\rho = -0.41$) and GHG emissions ($\rho = -0.15$) (Supplementary Table 3).

Sensitivity and subgroup analyses

Sensitivity analyses, only including participants without missing data, confirmed the direction and magnitude of our findings across human and environmental health outcomes (Supplementary Tables 6 and 7). Subgroup analyses for mortality and nutrient adequacy, and environmental outcomes, indicated only minor variations across sex, smoking status, body mass index and Mediterranean diet score (Supplementary Tables 10–17). Country-specific analyses showed that DSR_{Total} and DSR_{Plant} were consistently associated with lower all-cause mortality rates (Supplementary Table 8). Nevertheless, in contrast to the pooled analysis, DSR_{Plant} was positively associated with GHG emissions and land use in the Netherlands, Germany and Sweden (Supplementary Table 9). Models without adjustment for energy and food group intake yielded similar findings for DSR and nutrient adequacy and mortality, although the estimated environmental impacts were higher across DSR types (Supplementary Tables 18 and 19).

Discussion

This study, using data from the EPIC cohort, sought to explore the co-benefits of biodiverse diets for human (nutrient adequacy and mortality) and environmental health (dietary GHG emissions and

Table 1 | Baseline characteristics of 417,423 adults enrolled in the EPIC cohort across sex-specific quintiles (Q) of total dietary species richness (DSR_{Total})

	All (*n*=417,423)	Q1 (*n*=83,486)	Q2 (*n*=83,485)	Q3 (*n*=83,484)	Q4 (*n*=83,484)	Q5 (*n*=83,484)
DSR_{Total} (count of unique species consumed per year)	66.5 (15.5)	41.0 (7.49)	60.5 (3.51)	69.8 (2.18)	77.0 (2.58)	84.0 (3.02)
DSR_{Animal} (count of unique animal species consumed per year)	12.5 (4.29)	9.05 (2.95)	11.8 (2.77)	12.0 (5.30)	15.6 (3.86)	14.0 (2.87)
DSR_{Plant} (count of unique plant species consumed per year)	47.2 (12.1)	27.9 (6.19)	42.5 (3.50)	50.7 (4.67)	54.3 (3.27)	60.4 (3.60)
$PANDiet_{Overall}$ (%)	61.8 (8.30)	62.3 (8.39)	61.7 (7.55)	61.9 (7.56)	60.8 (7.94)	62.5 (9.75)
$PANDiet_{Micro}$ (%)	67.7 (15.8)	67.2 (16.1)	64.5 (17.3)	68.3 (14.7)	69.4 (14.9)	69.1 (15.3)
Dietary GHG emissions (usual $kgCO_2e$ per day)	5.21 (1.82)	5.12 (1.78)	4.92 (1.79)	5.12 (1.88)	5.54 (1.82)	5.33 (1.75)
Dietary land use (usual m^2 per day)	6.73 (2.60)	6.67 (2.58)	6.51 (2.55)	6.74 (2.75)	7.15 (2.64)	6.56 (2.42)
Age at recruitment (years)	51.4 (10.0)	52.9 (8.56)	51.0 (9.78)	49.3 (11.2)	52.6 (9.03)	51.1 (10.8)
Sex, *n* (%)						
Male	131,782 (31.6)	26,357 (31.6)	26,357 (31.6)	26,356 (31.6)	26,356 (31.6)	26,356 (31.6)
Female	285,641 (68.4)	57,129 (68.4)	57,128 (68.4)	57,128 (68.4)	57,128 (68.4)	57,128 (68.4)
Country, *n* (%)						
France	67,920 (16.3)	19,618 (23.5)	13,511 (16.2)	14,247 (17.1)	15,722 (18.8)	4,822 (5.8)
Italy	44,547 (10.7)	2,156 (2.6)	13,798 (16.5)	15,128 (18.1)	10,395 (12.5)	3,070 (3.7)
Spain	39,990 (9.6)	36,443 (43.7)	3,494 (4.2)	50 (0.1)	3 (0)	0 (0%)
UK	75,372 (18.1)	184 (0.2)	242 (0.3)	16,938 (20.3)	13,887 (16.6)	44,121 (52.8)
The Netherlands	36,538 (8.8)	2,194 (2.6)	15,544 (18.6)	15,452 (18.5)	3,348 (4.0)	0 (0)
Germany	49,352 (11.8)	224 (0.3)	1,765 (2.1)	4,235 (5.1)	13,757 (16.5)	29,371 (35.2)
Sweden	48,690 (11.7)	20,870 (25.0)	26,998 (32.3)	815 (1.0)	7 (0)	0 (0)
Denmark	55,014 (13.2)	1,797 (2.2)	8,133 (9.7)	16,619 (19.9)	26,365 (31.6)	2,100 (2.5)
Marital status, *n* (%)						
Single, divorced, separated or widowed	72,765 (17.4)	12,360 (14.8)	14,898 (17.8)	16,557 (19.8)	11,410 (13.7)	17,540 (21.0)
Married or living together	243,181 (58.3)	31,813 (38.1)	55,559 (66.5)	48,883 (58.6)	44,582 (53.4)	62,344 (74.7)
Unknown	101,477 (24.3)	39,313 (47.1)	13,028 (15.6)	18,044 (21.6)	27,492 (32.9)	3,600 (4.3)
Educational level, *n* (%)						
None or primary school completed	119,166 (28.5)	41,857 (50.1)	26,740 (32.0)	17,996 (21.6)	19,208 (23.0)	13,365 (16.0)
Technical/professional school	91,862 (22.0)	9,788 (11.7)	17,467 (20.9)	18,785 (22.5)	21,644 (25.9)	24,178 (29.0)
Secondary school	84,394 (20.2)	14,644 (17.5)	21,353 (25.6)	20,397 (24.4)	17,386 (20.8)	10,614 (12.7)
Longer education (including university degree)	10,518 (25.2)	15,846 (19.0)	16,984 (20.3)	23,484 (28.1)	22,552 (27.0)	26,252 (31.4)
Unknown	16,883 (4.0)	1,351 (1.6)	941 (1.1)	2,822 (3.4)	2,694 (3.2)	9,075 (10.9)
Smoking status, *n* (%)						
Never	208,205 (49.9)	43,065 (51.6)	42,255 (50.6)	40,536 (48.6)	39,775 (47.6)	42,574 (51.0)
Current	113,120 (27.1)	18,563 (22.2)	19,978 (23.9)	23,322 (27.9)	24,983 (29.9)	26,274 (31.5)
Former	89,562 (21.5)	20,282 (24.3)	20,223 (24.2)	18,657 (22.3)	17,580 (21.1)	12,820 (15.4)
Unknown	6,536 (1.6)	1,576 (1.9)	1,029 (1.2)	969 (1.2)	1,146 (1.4)	1,816 (2.2)
Physical activity, *n* (%)						
Inactive	87,030 (20.8)	25,343 (30.4)	18,240 (21.8)	12,796 (15.3)	12,729 (15.2)	17,922 (21.5)
Moderately inactive	143,126 (34.3)	29,471 (35.3)	27,211 (32.6)	27,789 (33.3)	28,957 (34.7)	29,698 (35.6)
Moderately active	101,116 (24.2)	17,909 (21.5)	19,876 (23.8)	20,990 (25.1)	22,129 (26.5)	20,212 (24.2)
Active	79,296 (19.0)	10,394 (12.5)	16,025 (19.2)	19,664 (23.6)	18,999 (22.8)	14,214 (17.0)
Unknown	6,855 (1.6)	369 (0.4)	2,133 (2.6)	2,245 (2.7)	670 (0.8)	1,438 (1.7)
Alcohol intake, *n*(%)						
Non-drinker	50,521 (12.1)	23,938 (28.7)	13,150 (15.8)	5,365 (6.4)	3,687 (4.4)	4,381 (5.2)
>0 to 6 g per day	118,264 (28.3)	16,633 (19.9)	30,880 (37.0)	25,407 (30.4)	20,652 (24.7)	24,692 (29.6)
>6 to 12 g per day	108,871 (26.1)	17,260 (20.7)	18,578 (22.3)	21,606 (25.9)	24,415 (29.2)	27,012 (32.4)

Table 1 (continued) | Baseline characteristics of 417,423 adults enrolled in the EPIC cohort across sex-specific quintiles (Q) of total dietary species richness (DSR_{Total})

	All (*n*=417,423)	Q1 (*n*=83,486)	Q2 (*n*=83,485)	Q3 (*n*=83,484)	Q4 (*n*=83,484)	Q5 (*n*=83,484)
>12 to 24 g per day	70,103 (16.8)	12,197 (14.6)	10,397 (12.5)	14,923 (17.9)	17,410 (20.9)	15,176 (18.2)
>24 g per day	69,664 (16.7)	13,58 (16.1)	10,480 (12.6)	16,183 (19.4)	17,320 (20.7)	12,223 (14.6)
Body mass index (kg m^{-2})	25.3 (4.21)	26.3 (4.54)	25.3 (4.16)	24.7 (4.01)	25.1 (4.06)	25.3 (4.12)
Energy intake (kcal per day)	2,110 (620)	2,080 (633)	2,050 (629)	2,130 (609)	2,190 (612)	2,100 (610)
Vegetables (g per day)	202 (131)	218 (144)	163 (121)	207 (134)	209 (120)	211 (128)
Fruit (g per day)	231 (180)	246 (198)	223 (174)	239 (181)	227 (168)	218 (177)
Cereal (g per day)	222 (115)	204 (110)	222 (118)	242 (120)	228 (113)	212 (111)
Total meat (g per day)	102 (62.4)	108 (63.5)	101 (56.9)	93.4 (69.0)	109 (63.1)	97.4 (57.4)
Processed meat (g per day)	32.9 (31.4)	36.7 (33.4)	32.6 (28.2)	25.5 (27.4)	31.7 (30.1)	37.7 (35.4)
Red meat (g per day)	44.3 (36.9)	39.7 (35.9)	44.4 (36.5)	48.6 (40.6)	52.7 (38.1)	36.1 (29.9)
Fish and shellfish (g per day)	33.6 (30.7)	46.4 (38.2)	26.5 (27.3)	23.2 (25.0)	37.6 (27.8)	34.6 (28.0)
Cakes and biscuits (g per day)	42.5 (44.1)	36.1 (41.4)	39.2 (38.0)	37.3 (38.3)	41.3 (44.2)	58.8 (52.9)
Condiments and sauces (g per day)	22.4 (20.6)	21.0 (21.9)	15.5 (17.0)	24.4 (21.3)	22.4 (18.2)	28.7 (21.9)
Mediterranean diet score (0–18 points), *n* (%)						
Low (0–6 points)	120,260 (28.8)	20,324 (24.3)	37,342 (44.7)	22,948 (27.5)	21,455 (25.7)	18,191 (21.8)
Medium (7–12 points)	224,556 (53.8)	43,922 (52.6)	35,744 (42.8)	44,731 (53.6)	48,468 (58.1)	51,691 (61.9)
High (13–18 points)	72,607 (17.4)	19,240 (23)	10,399 (12.5)	15,805 (18.9)	13,561 (16.2)	13,602 (16.3)

Continuous and ordinal variables are presented as the mean±s.d. and categorical variables as the frequency and proportion. DSR_{Total} includes the number of plant, fungal, animal and unidentified species. DSR_{Plant} includes the number of plant and fungal species. DSR_{Animal} includes the number of animal species. $PANDiet_{Overall}$ considers both macronutrients and micronutrients. $PANDiet_{Micro}$ considers only micronutrients. Physical activity categories (inactive, moderately inactive, moderately active and active) were derived from combining occupational physical activity with time spent in recreational activities. Educational levels were categorized according to the highest attained qualification (none or primary school, technical/professional school, secondary school and longer education, including university degrees).

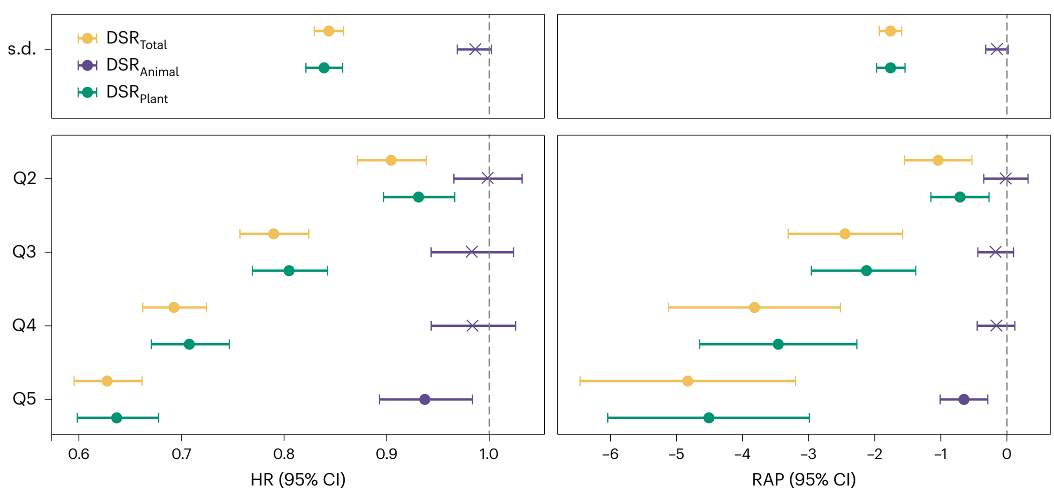

Fig. 1 | HRs and rate advancement periods and their 95% CIs for the relationship between DSR and all-cause mortality among 417,423 adults enrolled in the EPIC study. Cox regression models were stratified according to centre, age at recruitment (1-year intervals) and sex, and adjusted for marital status, educational level, smoking status, physical activity, alcohol intake, energy intake and consumption of vegetables, fruit, cereals, total meat, processed meat, red meat, fish and shellfish, cakes and biscuits and condiment and sauces. Individuals with unknown and missing data for marital status, educational level, smoking status, physical activity and alcohol intake were included in the analysis as a separate 'unknown' category. To ensure comparability, the various types of DSR were mean-standardized. Therefore, β coefficients reflect a 1-s.d. increment in the various types of DSR. For the analyses across quintiles, the first quintile served as the reference. DSR_{Total} includes the number of plant, animal and unidentified species. DSR_{Plant} includes the number of plant and fungal species. DSR_{Animal} includes the number of animal species. HRs with $P < 0.001$ are represented with dots and those with $P > 0.05$ are marked with crosses.

land use). Our analysis showed positive associations between all types of DSR and the probability of micronutrient adequacy, in parallel to inverse associations with mortality rates, with the exception of DSR_{Animal}. To clarify, DSR_{Animal} was inversely associated with the probability of overall nutrient adequacy but positively and neutrally associated with micronutrient adequacy and mortality rate, respectively. Findings for DSR_{Animal} and DSR_{Plant} provide additional insights into the relationships with mortality rates, as previous analyses solely assessed DSR_{Total}. In addition, DSR_{Total} did not show clear associations with environmental impacts. DSR_{Animal} was positively associated with GHG emissions and land use, while DSR_{Plant} showed inverse associations with these environmental impacts. These findings suggest the

Fig. 2 | β coefficients and 95% CIs of the linear relationship between DSR and nutrient adequacy and environmental impacts among 417,423 adults enrolled in the EPIC study. Linear models included centre as a random effect and were adjusted for age at recruitment (1-year intervals), sex, marital status, educational level, smoking status, physical activity, alcohol intake, energy intake and intakes of vegetables, fruit, cereals, total meat, processed meat, red meat, fish and shellfish, cakes and biscuits and condiments and sauces. Individuals with unknown and missing data for marital status, educational level, smoking status, physical activity and alcohol intake were included in the analysis as a separate 'unknown' category. To ensure comparability, the various types of DSR were mean-standardized. Therefore, β coefficients reflect a 1-s.d. increment in the various types of DSR. For the analyses across quintiles, the first quintile served as the reference. DSR_{Total} includes the number of plant, animal and unidentified species. DSR_{Plant} includes the number of plant and fungal species. DSR_{Animal} includes the number of animal species. $PANDiet_{Overall}$ considers both macronutrients and micronutrients. $PANDiet_{Micro}$ considers only micronutrients. Land use (m^2 per day).

relevance of diversifying human diets by increasing the number of plant species consumed.

The observed positive association between DSR and $PANDiet_{Micro}$ may be attributed to the 'sampling' or 'additive' effect, where higher DSR increases the chance of consuming micronutrient-rich species and a bounty of other health-protective metabolites. Furthermore, consuming a greater variety of species reduces the risk of excessive intake of specific nutrients, residues, contaminants and non-nutrients, thereby mitigating potential toxic effects[6]. DSR_{Animal} showed inverse associations for the PANDiet score encompassing macronutrients and micronutrients relative to the score including only micronutrients. This disparity can be attributed to the presence of a moderation subscore, calculated for nutrients for which the usual intake should not exceed a reference value, else assigning penalties. For example, one should not exceed a daily protein intake of 2.2 g kg^{-1} body weight[20]. Conversely, higher DSR could increase micronutrient adequacy by reducing the probability of exceeding these upper limits. When examining the different associations for DSR_{Plant} compared with DSR_{Animal} with nutrient adequacy, differences in nutrient profiles could be a contributor. Additionally, the PANDiet scores contain more vitamins than minerals, further contributing to the imbalance between plant and animal species. Moreover, the lower average number of animal species consumed could further accentuate this discrepancy.

Regarding environmental impacts, no associations were found between DSR_{Total} and dietary GHG emissions and land use. This shows how the metric of DSR is associated with reduced mortality rates and greater nutrient adequacy, without incurring additional GHG emissions or land use. Furthermore, a positive association was discerned between the DSR_{Animal} and dietary GHG emissions and land use. In the main analysis, an inverse association was found between DSR_{Plant} and dietary GHG emissions and land use. Unlike previous studies, which often focused on food groups or specific items, and primarily considered intake quantities, our study uniquely centred on the diversity of species consumed from whole diets, regardless of the quantities consumed. Additionally, while previous studies relied on estimations from model-based assumptions to quantify environmental impacts, we used data from life-cycle assessments of food and beverages, providing a more comprehensive and accurate estimation[21]. However, associations between DSR_{Plant} and dietary GHG emissions and land use were inconsistent in the country-specific analyses. These inconsistencies may stem from several causes. First, methodological differences, such as questionnaire design and the number of items included in the food list, might have influenced results. Second, environmental impacts vary greatly according to plant species, with some exerting higher impacts than others. By contrast, DSR_{Animal} impacts may exhibit less variability because of the narrower range of species consumed and their relatively consistently high environmental impacts. Third, individuals consuming a higher diversity of plant species may also be more likely to consume lower-impact animal species, and vice versa, potentially influencing the observed associations. Fourth, GHG emissions and land use were calculated for food items, and inconsistent associations in the Netherlands, Germany and Sweden might be linked to other differences in dietary patterns, such as quantities consumed, not captured by DSR[22].

DSR represents a conceptual advancement in understanding dietary diversity because it quantifies diversity both between and within food groups. By examining the number of distinct biological species within a diet, DSR offers a more granular and cross-cutting

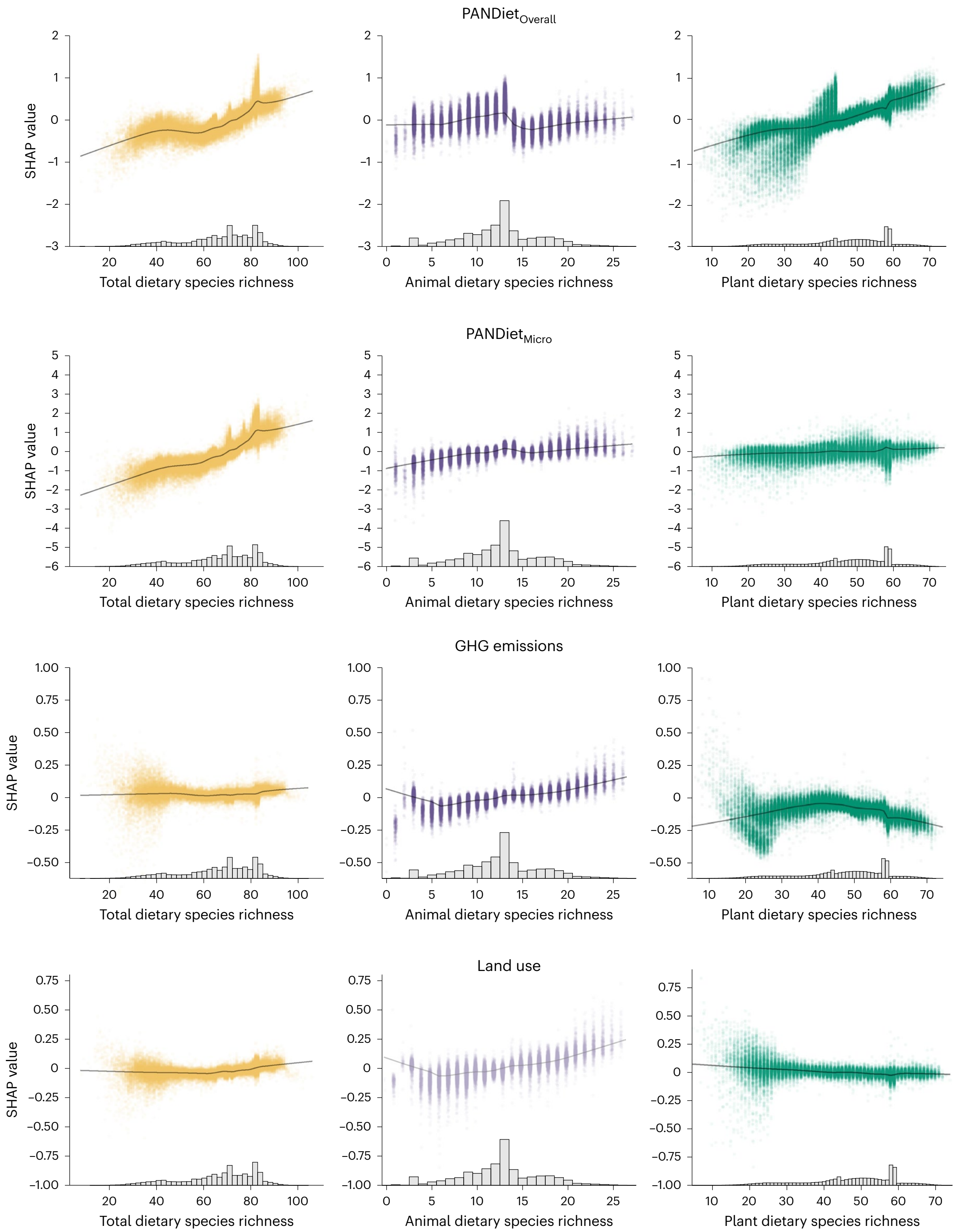

Fig. 3 | Correlation between DSR and SHAP values of nutrient adequacy and environmental impacts among 417,423 adults enrolled in the EPIC study. SHAP values quantify the contribution of each feature to the model's prediction. Higher SHAP values indicate greater influence of a variable on the outcome of interest. Locally weighted scatterplot smoothing (LOWESS) curves were applied to visualize the relationship between DSR and the respective SHAP values, offering insights into the trend of the data over the full range of DSR values. Models were adjusted for centre, age at recruitment (1-year intervals), sex, marital status, educational level, smoking status, physical activity, alcohol intake, energy intake and intakes of vegetables, fruit, cereals, total meat, processed meat, red meat, fish and shellfish, cakes and biscuits and condiments and sauces. Individuals with unknown and missing data for marital status, educational level, smoking status, physical activity and alcohol intake were included in the analysis as a separate 'unknown' category. DSR$_{Total}$ includes the number of plant, animal and unidentified species. DSR$_{Plant}$ includes the number of plant and fungal species. DSR$_{Animal}$ includes the number of animal species. PANDiet$_{Overall}$ considers both macronutrients and micronutrients. PANDiet$_{Micro}$ considers only micronutrients. All *P* values of the Spearman's rank correlation coefficients were <0.001.

assessment of dietary diversity. Importantly, DSR was not designed as a metric of overall diet quality or an environmental assessment tool; rather, it aims to capture food biodiversity, which might complement other known characteristics of sustainable healthy diets and may lead to the preservation of finite genetic resources in diets and on farms. The modest associations observed with nutrient adequacy and environmental impacts reflect this distinction. Mediation analysis suggests that nutrient adequacy explained only a small proportion of the association between DSR and mortality rates, suggesting that dietary diversity and nutrient adequacy represent two distinct characteristics of a healthy diet. The stronger, more consistent associations with mortality suggest that biodiverse diets confer benefits beyond more adequate nutrient intakes.

This study has several strengths. First, our analyses included both human and environmental health outcomes. Second, the prospective design within the EPIC cohort, with over 400,000 individuals and extensive follow-up, allows for stratification and enhancing generalizability to similar contexts. However, it is important to refrain from extrapolating these results to the entire European population, given the study's focus on middle-aged individuals with generally more health-conscious behaviour. Third, the standardized dietary assessments covered approximately 250 unique species and considered the contribution of each food and beverage item to GHG emissions and land use. Although this study solely focused on GHG emissions and land use to assess environmental impacts, adopting a biodiverse diet may yield additional environmental benefits by supporting biodiversity and fostering more resilient ecosystems that are better adapted to shocks[10,11]. Additionally, the study's holistic approach aligns with Sustainable Development Goals, providing valuable insights into DSR's utility and relevance. The metric of DSR analyses food items based on their specific species of origin rather than comparing one's diet with a reference such as the EAT–Lancet planetary health diet. However, DSR could work in complementary fashion to this EAT–Lancet diet, already suggesting the existence of co-benefits[19]. By distinguishing between animal and plant species, the study offers a clearer understanding and facilitates the formulation of guidelines concerning their impacts on both dietary and environmental health. Lastly, using machine learning models with SHAP dependence plots, in addition to linear mixed models, enhanced the robustness of our results[23].

In this study, several limitations must be acknowledged. First, the estimation of DSR involved several methodological challenges. Self-reported dietary intakes do not accurately capture consumption patterns; therefore, infrequently consumed foods may be underestimated or omitted during questionnaire completion. Furthermore, differences in dietary assessment methods and the number of items included between centres limited consistent taxonomic classification. Insufficient taxonomic detail also hampers the subdivision of species into cultivars with differing nutrient content, and some foods entirely lack species-level information, probably leading to an underestimation of DSR[6]. These limitations for DSR are compounded by general dietary assessment constraints. Food frequency questionnaires are usually less precise for calculating absolute dietary intakes, while better suited for ranking individuals according to their intakes. Subgroup analyses according to country aimed to account for these differences in dietary assessment. Our study also used a single assessment of self-reported dietary intakes at baseline, which might not accurately capture dietary patterns over the long follow-up period. While diets are dynamic and subject to change over time, it is generally assumed that these estimations reflect typical eating behaviours throughout middle-aged adult life[24]. Regarding outcome assessments, the PANDiet score used estimated average requirements and adequate intakes as reference values. This approach assumes a normal distribution of nutrient requirements, which does not hold true for certain nutrients like iron, potentially affecting the accuracy of nutritional adequacy assessments for these specific nutrients. Regarding environmental impacts, dietary GHG emission and land use data were derived from the SHARP-Indicators Database, providing standardized but non-country-specific estimates. While this approach captures pan-European impacts, it does not account for country-specific variations. This limitation may have led to outcome misclassification, albeit non-differentially given the study focus on overall associations rather than individual estimates. Future studies would benefit from incorporating country-specific environmental impact estimates to enhance precision and potentially reveal regional variations in these associations[25]. Finally, residual confounding might be present as this study did not address broader contextual factors, including social, ethical, economic, cultural and food safety considerations, which could also influence dietary choices and affect the accessibility, affordability and acceptability of consuming diverse species. Future research would benefit from exploring beta diversity (the dissimilarity of species composition in diets among individuals and across countries) to provide deeper insights into these multidimensional characteristics of dietary patterns.

In summary, our study suggests that transitioning towards more biodiverse diets is associated with lower mortality rates, which is not mediated through nutrient adequacy, with plant species showing stronger benefits when compared with animal species. Specifically, diversifying the intake of plant species might be instrumental in achieving co-benefits for planetary health, as highlighted by the inverse associations with GHG emissions and land use. Nonetheless, the mechanisms underlying these associations remain incompletely understood. Further research should explore the pathways through which biodiverse diets influence mortality rates and examine the role of other dietary characteristics across a wide range of contexts.

Methods

Inclusion and ethics

The present study is relevant to guide local agrifood system transformation and nutrition research. Local research teams were invited to contribute to the paper and were included as co-authors in accordance with ICMJE criteria.

The EPIC study was approved by local ethics committees and by the internal review board of the International Agency for Research on Cancer (IARC).

The EPIC cohort

EPIC is a large multicentre cohort study examining the interplay between metabolic, lifestyle and environmental factors related to cancer and chronic diseases. Detailed information on the EPIC cohort is available in Supplementary Text 1. Briefly, over 500,000 individuals aged 25–70 were recruited between 1992 and 2000 across 23 centres in 10 European countries: Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden and the UK. Dietary intake was assessed at enrolment using validated country-specific or centre-specific dietary questionnaires designed to capture habitual consumption over the previous 12 months[26]. Vital status and date of death were determined through record linkage with cancer registries, boards of health, health insurance registries, pathology registries or through active follow-up and next of kin. Exclusions were made for missing lifestyle or dietary data, extreme energy intakes, lack of follow-up and prevalent diseases at baseline. The EPIC-Greece and EPIC-Norway cohorts were not available for this analysis. Our analysis included 417,423 participants, with 45,489 deaths recorded from 1992 to 2014 (Supplementary Fig. 1).

Food biodiversity calculation

Food biodiversity was quantified using the DSR, as described previously[6]. Despite the existence of several metrics for quantifying food biodiversity, DSR emerges as a particularly robust measure because of its intuitive interpretation, demonstrated construct validity and capacity for meaningful cross-context comparisons[14]. To calculate an individual's food biodiversity, the total count of distinct biological species present in each food, beverage and recipe was tallied. For all countries, composite dishes were broken down into their individual ingredients using standard recipes. However, ingredients like herbs and spices, used in small amounts, could bias the DSR calculation. To address this, three DSR scenarios were assessed previously: (1) DSR, including all food items from the EPIC food list (including ingredients from standard recipes, regardless of quantity); (2) DSR excluding the lowest 5% of species intake (g per day) from each food group; and (3) DSR excluding the lowest 10% of species intake (g per day) from each food group. These scenarios did not show any significant differences, so all food items were used for the DSR calculation[6]. Food items consumed 'never or less than once per month' were recalled under one category and these species were not considered for the DSR. Moreover, quantities (g per day) were disregarded for the DSR computation because our interest was the sum of distinct species consumed per year. The DSR was expressed as the number of species consumed per person per year (DSR_{Total}) and further disaggregated into animal (DSR_{Animal}) and plant (DSR_{Plant}) species, with DSR_{Plant} including both plant and fungal species.

Nutrient adequacy

To evaluate an individual's probability of nutrient adequacy, we calculated their PANDiet score. The PANDiet score is based on 26 nutrients and was derived using estimated average requirements, or when unavailable, adequate intake levels provided by the European Food Safety Authority[20]. PANDiet scores, which ranged between 0% and 100%, were calculated for both macronutrients and micronutrients ($PANDiet_{Overall}$), and for micronutrients only ($PANDiet_{Micro}$) (Supplementary Text 2 and Supplementary Table 1).

Environmental impacts

Dietary GHG emissions and land use were estimated using the SHARP-Indicators Database, which leverages life-cycle analysis estimates based on the environmental impacts of food production, packaging, transport and home preparation[22]. Food items were matched between the EPIC and SHARP-Indicators Databases using their FoodEx 2 base terms based on the European Food Safety Authority food description and classification system[27]. GHG emissions were expressed as kilogram of CO_2equivalents per kg of food per day and land use as metre squared per year per kg of food per day.

Statistical analysis

To quantify the potential associations between DSR and mortality, nutrient adequacy and dietary environmental impacts, multivariable-adjusted Cox and linear regression models were fitted. DSR was considered both as a continuous exposure and categorized into sex-specific quintiles. Models were fitted for each type of DSR. Multivariable adjustments were performed using a 10% change-in-estimate criterion for β coefficients. Thereafter, directed acyclic graphs were used to ensure that potential collider bias was avoided.

This criterion was used on different food groups and sociodemographic variables given the limited knowledge of factors related to DSR and our outcomes of interest. To ensure consistency, confounding variables meeting the criterion in any model were included in all analyses. Study centre was included in every model, to address differences in questionnaire design, follow-up procedures and other centre-specific effects. Based on the criterion, confounders included were age at recruitment (1-year intervals), sex (male or female), marital status (single, divorced, separated or widowed; married or living together; unknown), educational level (none or primary school; technical or technical degree; secondary; higher education; unknown), smoking status (never; current; former; unknown), physical activity (inactive; moderately inactive; moderately active; active; unknown), alcohol intake (non-drinker; >0 to 6 g per day; >6 to 12 g per day; >12 to 24 g per day; >24 g per day), energy intake (kcal per day) and the consumption of vegetables, fruit, cereals, total meat, processed meat, red meat, fish and shellfish, cakes and biscuits and condiment and sauces, all in g per day. To evaluate the possibility of multicollinearity, variance inflation factors were used; all yielded values below 3.

Previous analyses of DSR and all-cause and cause-specific mortality only used DSR_{Total}[6,14]. Because of differences in sample size and expansion to include DSR_{Animal} and DSR_{Plant} in this study, the analysis was reconducted using multivariable-adjusted Cox regression models. HRs and their corresponding 95% CIs were computed for DSR, both as continuous variables and sex-specific quintiles. Time until death or loss to, or end of follow-up, whichever came first, served as the underlying time variable. Models were stratified according to sex, age at recruitment (1-year intervals) and centre, and adjusted for the identified confounders. Shoenfeld residuals, according to follow-up time (years), confirmed the Cox proportional hazards assumptions. Associations between DSR and mortality rates were quantified using RAP[28]. A 1-s.d. increment in DSR was examined, along with four scenarios, comparing individuals in the second, third, fourth and fifth quintiles with those in the first. RAP estimates represent the impact of a given exposure on the risk of death by determining the expected time (in years) by which the risk of death is accelerated or decelerated for exposed individuals compared with non-exposed individuals. RAP was calculated by dividing the log(HR) estimates by the log of the coefficient of age at recruitment.

Associations between DSR and nutrient adequacy and dietary GHG emissions and land use were assessed using multivariable-adjusted linear regression models with study centre as a random effect and all other confounders as fixed effects. Machine learning models complemented the linear models to explore potential non-linear relationships. XGBoost, a gradient-boosting algorithm known for its efficacy, was selected based on its demonstrated applicability, as it can handle non-linear data, capture interactions, is robust to outliers and shows high performance and speed[29]. Hyperparameters were optimized using Bayesian optimization[30] (Supplementary Table 2). To explore the impacts of DSR on predictions, the feature importance of DSR, in other words, how much DSR contributes to the prediction process, was evaluated. Feature analysis was conducted using SHAP, a method capable of elucidating the importance of a feature on output predictions of machine learning models[23]. SHAP values measure the contribution of each feature to the overall prediction. SHAP values provide a clearer understanding of the relationship between DSR and the predicted outcomes by quantifying how much DSR influences the model's output, whether positively or negatively. DSR was plotted against SHAP values to visually illustrate this relationship. LOWESS curves were fitted to the SHAP values to capture trends in how DSR influences predictions. Additionally, Spearman's rank correlations between DSR and human and environmental health outcomes were assessed.

To examine the role of nutrient adequacy in the association between DSR and all-cause mortality, we conducted mediation analyses. We first confirmed that DSR predicted mortality and verified that DSR predicted the mediator, namely the PANDiet scores; finally, we examined whether DSR remained associated with mortality after adjusting for the PANDiet scores. The mediation analyses were adjusted for the same confounders as our main analyses. We estimated the pure natural direct effect (PNDE), total natural indirect effect (TNIE), total effect (TE) and proportion mediated (PM) ($PM = 100 \times (PNDE \times (TNIE - 1))/(TE - 1)$) using the counterfactual framework, with 95% CIs derived through 1,000 bootstrap iterations.

In the main analysis, missing data of categorical variables were treated as a separate 'unknown' category. Sensitivity analyses were conducted to assess the overall robustness of our models, using a complete case analysis excluding participants with this 'unknown' category. In addition, subgroup analyses were performed for sex, country, body mass index, smoking status and diet quality using the Mediterranean diet score to assess heterogeneity[31]. Additionally, the analysis was run without adjustment for energy intake and food group intake. Statistical tests were two-sided and $P < 0.05$ was considered statistically significant. All analyses were performed using Python v.3.9.13 and R v.4.04.1.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

EPIC data and biospecimens are available to investigators in the context of research projects that are consistent with the legal and ethical standard practices of IARC/World Health Organization (WHO) and the EPIC centres. The use of a random sample of anonymized data from the EPIC study can be requested by contacting https://epic.iarc.fr/. For information on the EPIC data access policy and on how to submit an application for gaining access to EPIC data and/or biospecimens, please follow the instructions at iarc.who.int.

Code availability

Analytical code is available via Code Ocean at https://codeocean.com/capsule/0563231/tree.

References

1. Tilman, D. & Clark, M. Global diets link environmental sustainability and human health. *Nature* **515**, 518–522 (2014).
2. Benton, T. G., Bieg, C., Harwatt, H., Pudasaini, R. & Wellesley, L. *Food System Impacts on Biodiversity Loss Three Levers for Food System Transformation in Support of Nature* (Chatham House, 2021).
3. Willett, W. et al. Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. *Lancet* **393**, 447–492 (2019).
4. Beal, T., Massiot, E., Arsenault, J. E., Smith, M. R. & Hijmans, R. J. Global trends in dietary micronutrient supplies and estimated prevalence of inadequate intakes. *PLoS ONE* **12**, e0175554 (2017).
5. Afshin, A. et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **393**, 1958–1972 (2019).
6. Hanley-Cook, G. T. et al. Food biodiversity and total and cause-specific mortality in 9 European countries: an analysis of a prospective cohort study. *PLoS Med.* **18**, e1003834 (2021).
7. Khoury, C. K. et al. Increasing homogeneity in global food supplies and the implications for food security. *Proc. Natl Acad. Sci. USA* **111**, 4001–4006 (2014).
8. Kyttä, V., Hyvönen, T. & Saarinen, M. Land-use-driven biodiversity impacts of diets—a comparison of two assessment methods in a Finnish case study. *Int. J. Life Cycle Assess.* **28**, 1104–1116 (2023).
9. Campbell, B. M. et al. Agriculture production as a major driver of the Earth system exceeding planetary boundaries. *Ecol. Soc.* **22**, 8 (2017).
10. Folke, C. et al. Regime shifts, resilience, and biodiversity in ecosystem management. *Annu. Rev. Ecol. Evol. Syst.* **35**, 557–581 (2004).
11. Oliver, T. H. et al. Biodiversity and resilience of ecosystem functions. *Trends Ecol. Evol.* **30**, 673–684 (2015).
12. Kennedy, G., Lee, W. T. K., Termote, C., Charrondière, R. & Tung, J. Y. A. *Guidelines on Assessing Biodiverse Foods in Dietary Intake Surveys* (FAO, 2017).
13. Verger, E. O. et al. Dietary diversity indicators and their associations with dietary adequacy and health outcomes: a systematic scoping review. *Adv. Nutr.* **12**, 1659–1672 (2021).
14. Hanley-Cook, G. et al. Dietary species richness provides a comparable marker for better nutrition and health across contexts. *Nat. Food* **6**, 577–586 (2025).
15. Lachat, C. et al. Dietary species richness as a measure of food biodiversity and nutritional quality of diets. *Proc. Natl Acad. Sci. USA* **115**, 127–132 (2018).
16. Huybrechts, I. et al. Food biodiversity and gastrointestinal cancer risk in nine European countries: analysis within a prospective cohort study. *Eur. J. Cancer* **210**, 114258 (2024).
17. Springmann, M. et al. Options for keeping the food system within environmental limits. *Nature* **562**, 519–525 (2018).
18. Sorooshian, S. The sustainable development goals of the United Nations: a comparative midterm research review. *J. Clean. Prod.* **453**, 142272 (2024).
19. Laine, J. E. et al. Co-benefits from sustainable dietary shifts for population and environmental health: an assessment from a large European cohort study. *Lancet Planet Health* **5**, e786–e796 (2021).
20. Verger, E. O., Mariotti, F., Holmes, B. A., Paineau, D. & Huneau, J.-F. Evaluation of a diet quality index based on the probability of adequate nutrient intake (PANDiet) using national French and US dietary surveys. *PLoS ONE* **7**, e42155 (2012).
21. Quam, V., Rocklöv, J., Quam, M. B. M. & Lucas, R. Assessing greenhouse gas emissions and health co-benefits: a structured review of lifestyle-related climate change mitigation strategies. *Int. J. Environ. Res. Public Health* **14**, 468 (2017).
22. Mertens, E. et al. SHARP-Indicators Database towards a public database for environmental sustainability. *Data Brief* **27**, 104617 (2019).
23. Lundberg, S. M. & Lee, S. A unified approach to interpreting model predictions. Preprint at https://arxiv.org/abs/1705.07874 (2017).
24. Willett, W. *Nutritional Epidemiology* (Oxford Univ. Press, 2012).
25. Poore, J. & Nemecek, T. Reducing food's environmental impacts through producers and consumers. *Science* **360**, 987–992 (2018).
26. Huybrechts, I. et al. Characterization of the degree of food processing in the European Prospective Investigation into Cancer and Nutrition: application of the Nova classification and validation using selected biomarkers of food processing. *Front. Nutr.* **9**, 1035580 (2022).
27. European Food Safety Authority (EFSA). The food classification and description system FoodEx 2 (revision 2). *EFSA* **12**, 804E (2015).
28. Brenner, H., Gefeller, O. & Greenland, S. Risk and rate advancement periods as measures of exposure impact on the occurrence of chronic diseases. *Epidemiology* **4**, 229–236 (1993).
29. Bentéjac, C., Csörgő, A. & Martínez-Muñoz, G. A comparative analysis of gradient boosting algorithms. *Artif. Intell. Rev.* **54**, 1937–1967 (2021).
30. Yang, L. & Shami, A. On hyperparameter optimization of machine learning algorithms: theory and practice. *Neurocomputing* **415**, 295–316 (2020).
31. Buckland, G. et al. Adherence to the Mediterranean diet and risk of coronary heart disease in the Spanish EPIC cohort study. *Am. J. Epidemiol.* **170**, 1518–1529 (2009).

Acknowledgements

The coordination of the European Investigation into Cancer and Nutrition (EPIC) study is financially supported by International Agency for Research on Cancer (IARC) and also by the Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, which has additional infrastructure support provided by the National Institute for Health and Care Research (NIHR) Imperial Biomedical Research Centre (BRC). The national cohorts are supported by the Danish Cancer Society (Denmark); the Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale and Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, the German Cancer Research Center (DKFZ), the German Institute of Human Nutrition

Potsdam-Rehbruecke (DIfE) and the Federal Ministry of Education and Research (BMBF) (Germany); the Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy and the Compagnia di SanPaolo and National Research Council (Italy); the Dutch Ministry of Public Health, Welfare and Sports (VWS), the Netherlands Cancer Registry (NKR), the LK Research Funds, the Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), the World Cancer Research Fund (WCRF) and Statistics Netherlands (the Netherlands); Health Research Fund (FIS)-Instituto de Salud Carlos III (ISCIII), the Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, and the Catalan Institute of Oncology (ICO) (Spain); the Swedish Cancer Society, the Swedish Research Council and the County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C8221/A29017 to EPIC-Oxford) and the Medical Research Council (1000143 to EPIC-Norfolk; MR/M012190/1 to EPIC-Oxford) (UK); and UiT The Arctic University of Norway (Norway). Funding for grant no. IIG_FULL_2020_034 was obtained from Wereld Kanker Onderzoek Fonds (WKOF) as part of the World Cancer Research Fund International grant programme. P.V. is supported by the NIHR Global Health Research Centre on Non-Communicable Diseases (NCDs) and Environmental Change (NIHR203247). E.K.C. was supported by the Ministry of National Education, Türkiye as part of their YLSY International Graduate Education Scholarship programme. Researchers were independent from the funders. The funders had no role in the collection, analysis and interpretation of the data, the writing of the manuscript, or the decision to submit the article for publication. Where authors are identified as personnel of IARC/WHO, the authors alone are responsible for the views expressed in this article, and they do not necessarily represent the decisions, policy or views of IARC/WHO.

Author contributions

Conceptualization: J. Berden, C.L. and I.H. Data curation: I.H., G.N. and P.F. Formal analysis: J. Berden, I.H. and G.N. Funding acquisition: G.T.H.-C., I.H., C.L. and M.T. Investigation: J. Berden, B.C., G.T.H.-C., E.K.C., P.V., G.N., G. Skeie, B. Srour, E.K.-G., M.T., J. Baudry, M.D.-T., J. Berden, Y.v.d.S., K.M., F.J., A.T., C.K., C.C.D., D.B.I., C.L.C., M.B.S., L.M., C.M., E.W., A.H., G. Severi, C.C., A.J.-Z., B. Sodano, C.C.-E., J.C., K.T., M.-D.C., M.-J.S., P.C., S.P., V.K., M.G., P.F., C.L. and I.H. Methodology: J. Berden, C.L. and I.H. Project administration: J. Berden, C.L. and I.H. Resources: G.T.H.-C., P.V., G.N., G. Skeie, Y.v.d.S., K.M., F.J., A.T., C.K., C.C.D., D.B.I., C.L.C., M.B.S., C.M., E.W., A.H., G. Severi, C.C., A.J.-Z., B. Sodano, C.C.-E., J.C., K.T., M.-D.C., M.-J.S., P.C., S.P., V.K., M.G., P.F., C.L. and I.H. Software: J. Berden. Supervision: C.L. and I.H. Visualization: J. Berden. Validation: G.N., C.L. and I.H. Writing—original draft: J. Berden. Writing—review and editing: B.C., G.T.H.-C., E.K.C., P.V., G.N., G. Skeie, B. Srour, E.K.-G., M.T., J. Baudry, M.D.-T., J. Berlivet, Y.v.d.S., K.M., F.J., A.T., C.K., C.C.D., D.B.I., C.L.C., M.B.S., L.M., C.M., E.W., A.H., G. Severi, C.C., A.J.-Z., B. Sodano, C.C.-E., J.C., K.T., M.-D.C., M.-J.S., P.C., S.P., V.K., M.G., P.F., C.L. and I.H.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43016-025-01214-y.

Correspondence and requests for materials should be addressed to Jeroen Berden.

Peer review information *Nature Food* thanks Maijaliisa Erkkol and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Jeroen Berden[1,2]✉, Bernadette Chimera[1], Giles T. Hanley-Cook[2], Emine Koc Cakmak[3], Paolo Vineis[4], Genevieve Nicolas[1], Guri Skeie[1,5], Bernard Srour[6], Emmanuelle Kesse-Guyot[6], Mathilde Touvier[6], Julia Baudry[6], Mélanie Deschasaux-Tanguy[6], Justine Berlivet[6], Yvonne van der Schouw[7], Kris Murray[8], Franziska Jannasch[9], Anne Tjønneland[10,11], Cecilie Kyrø[10], Christina C. Dahm[12], Daniel Borch Ibsen[12], Charlotte Le Cornet[13], Matthias B. Schulze[9,14], Lorenzo Mangone[4], Chloé Marques[15], Elisabete Weiderpass[16], Alicia Heath[4], Gianluca Severi[15,17], Claire Cadeau[15], Ana Jiménez-Zabala[18,19], Barbara Sodano[17,20], Carlota Castro-Espin[21,22], Jesús Castilla[23,24], Kostas Tsilidis[3,23], María-Dolores Chirlaque[24,25], Maria-Jose Sánchez[25,26,27], Paolo Contiero[28], Salvatore Panico[29], Verena Katzke[13], Marc Gunter[4], Pietro Ferrari[1], Carl Lachat[2,30] & Inge Huybrechts[1,30]

[1]Nutrition and Metabolism Branch, International Agency for Research on Cancer, Lyon, France. [2]Department of Food Technology, Safety and Health, Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium. [3]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK. [4]Cancer Epidemiology and Prevention Research Unit, School of Public Health, Imperial College London, London, UK. [5]Department of Community Medicine, UiT The Arctic University of Norway, Tromsø, Norway. [6]Université Sorbonne Paris Nord and Université Paris Cité, INSERM, INRAE, CNAM, Centre of Research in Epidemiology and StatisticS (CRESS), Nutritional Epidemiology Research Team (EREN), Bobigny, France. [7]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands. [8]Centre on Climate Change and Planetary Health, MRC Unit The Gambia at London School of Hygiene and Tropical Medicine, Fajara, The Gambia. [9]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany. [10]Danish Cancer Institute, Copenhagen, Denmark. [11]Department of Public Health, University of Copenhagen, Copenhagen, Denmark. [12]Department of Public Health, Aarhus University, Aarhus, Denmark. [13]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany. [14]Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany. [15]Paris-Saclay University, UVSQ, INSERM, Gustave Roussy, CESP, Villejuif, France. [16]International Agency for Research on Cancer, Lyon, France. [17]Department of Statistics, Computer Science, Applications 'G. Parenti', University of Florence, Florence, Italy.

[18]Ministry of Health of the Basque Government, Sub-Directorate for Public Health and Addictions of Gipuzkoa, San Sebastián, Spain. [19]Biogipuzkoa Health Research Institute, Group of Epidemiology of Chronic and Communicable Diseases, San Sebastián, Spain. [20]Centre for Biostatistics, Epidemiology, and Public Health, Department of Clinical and Biological Sciences, University of Turin, Orbassano, Italy. [21]Unit of Nutrition and Cancer, Catalan Institute of Oncology-ICO, L'Hospitalet de Llobregat, Barcelona, Spain. [22]Nutrition and Cancer Group, Epidemiology, Public Health, Cancer Prevention and Palliative Care Program, Bellvitge Biomedical Research Institute-IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain. [23]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece. [24]Department of Epidemiology, Regional Health Council, IMIB-Arrixaca, Murcia University, Murcia, Spain. [25]CIBER in Epidemiology and Public Health (CIBERESP), Madrid, Spain. [26]Escuela Andaluza de Salud Pública (EASP), Granada, Spain. [27]Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, Spain. [28]Environmental Epidemiology Unit, Fondazione IRCCS Istituto Nazionale dei Tumori di Milano, Milan, Italy. [29]School of Medicine, Federico II University, Naples, Italy. [30]These authors jointly supervised this work: Carl Lachat, Inge Huybrechts. ✉e-mail: jeroen.berden@ugent.be

Corresponding author(s): Jeroen Berden
NATFOOD-24102043A

Last updated by author(s): Jun 3, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Dietary intake data in EPIC was collected using EPIC soft (see Voss et al. EPIC-SOFT a European computer program for 24-hour dietary protocols Z Ernahrungswiss 1998 Sep;37(3):227-33. doi: 10.1007/s003940050021.
Data analysis	Analyses were performed using Python version 3.9.13 and R version 4.04.1. Python was used for the gradient boosting models annd data visualisation, while R was used for all other descriptive, and statistical analysis and data visualisation.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

EPIC data and biospecimens are available to investigators in the context of research projects that are consistent with the legal and ethical standard practices of

IARC/WHO and the EPIC Centers. The use of a random sample of anonymized data from the EPIC study can be requisted by contacting epic@iarc.fr. For information on EPIC data access policy and how to submit an application for gaining access to EPIC data and/or biospecimens, please follow the instructions at http://epic.iarc.fr/access/index.php.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Nutrient requirements are influenced by sex and sex-based analysis were performed for all different analysis assessed and can be found in the supplementary materials.
Reporting on race, ethnicity, or other socially relevant groupings	No variables for race or ethnicity were used in this study.
Population characteristics	The European Prospective Investigation into Cancer and Nutrition (EPIC) cohort includes a broad age range, predominantly middle-aged adults, with participants from various socioeconomic backgrounds and regions in Europe. It captures detailed dietary information, physical activity levels, body measurements, and lifestyle factors such as smoking and alcohol use, allowing for comprehensive assessments of health and disease patterns across different European populations.
Recruitment	The European Prospective Investigation into Cancer and Nutrition (EPIC) is a large multicentre cohort study examining the interplay between metabolic, lifestyle, and environmental factors related to cancer and chronic diseases. Over 500 000 individuals aged 25-70 were recruited between 1992 and 2000 across 23 centres in 10 European countries: Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden, and the United Kingdom.
Ethics oversight	EPIC data and biospecimens are available to investigators in the context of research projects that are consistent with the legal and ethical standard practices of IARC/WHO and the EPIC Centers. The use of a random sample of anonymized data from the EPIC study can be requested by contacting epic@iarc.fr. For information on EPIC data access policy and on how to submit an application for gaining access to EPIC data and/or biospecimens, please follow the instructions at http://epic.iarc.fr/access/index.php.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This manuscript uses data from a prospective cohort study (EPIC).
Research sample	The European Prospective Investigation into Cancer and Nutrition (EPIC) study and the sub-samples of enrolled adults (n=417,423) from Denmark (n=55,014), France (n=67,920), Germany (n=49,352), Italy (n=44,547), the Netherlands (n=36,538), Spain (n=39,990), Sweden (n=48,690), and the United Kingdom (n=75,372) included in the present analyses have been extensively described previously. Detailed information on the EPIC cohort is available in Supplemental text 1.
Sampling strategy	EPIC is a prospective cohort study that recruited participants from the Riboli E, Hunt K, Slimani N, Ferrari P, Norat T, Fahey M, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr. 2002;5:1113–24. pmid:12639222 .
Data collection	The European Prospective Investigation into Cancer and Nutrition (EPIC) study and the sub-samples of enrolled adults (n=417,423) from Denmark (n=55,014), France (n=67,920), Germany (n=49,352), Italy (n=44,547), the Netherlands (n=36,538), Spain (n=39,990), Sweden (n=48,690), and the United Kingdom (n=75,372) included in the present analyses have been extensively described previously. Detailed information on the EPIC cohort is available in Supplementary text 1
Timing	Participants from EPIC were enrolled from between 1992 and 2000 at each of the different centres. Dietary assessment was conducted during this period.
Data exclusions	In EPIC, we excluded participants with missing dietary information, those with an extreme ratio of energy intake to energy requirement (top and bottom 1%, as these values were considered physiologically implausible (71)), volunteers with null follow-up, those with prevalent disease at baseline (history of cancer, cardiovascular diseases, and diabetes), and all participants from the EPIC-Greece and EPIC-Norway cohort, due to administrative constraints.

Non-participation	Following the regulations for medical research on humans, data was only collected from participants that agreed to take part in the study.
Randomization	Not Applicable.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	/
Novel plant genotypes	/
Authentication	/