Manuscript: "Reproducibility of *a posteriori* dietary patterns across time and studies: a scoping review"

Valeria Edefonti

Online Supplementary Material

Supplemental materials and methods

In **Supplemental Table 1**, we introduce the basic terminology we adopted in the current review, as well as the statistical tools used for their assessment. In **Supplemental Figure 1** we introduce prototypical paths of DP identification processes related to reproducibility and validity of DPs. Dietary patterns are identifiable within any study design and starting from any dietary assessment tool source. If one dietary source is used at one time point, the assessment of DP reproducibility arises from the use of different statistical approaches for DP identification [Panel (A)]. Within the validation study of a new food-frequency questionnaire (FFQ), the same FFQ was administered twice (within 1 year) and compared with a gold standard dietary assessment tool [a dairy record (DR) or (multiple administration of) a 24-hour recall (24HR)] carried out on the same time interval and sample; DP reproducibility is assessed comparing the 2 sets of FFQ-based DPs, whereas relative validity of DPs is assessed comparing FFQ-based and gold-standard-based DPs [Panel (B)]. When either cohort studies or multiple waves of the same survey are available, a dietary assessment tool is administered to the same subjects in multiple occasions over longer time periods and the comparison of sets of DPs at the available measurement occasions allows for the evaluation of stability of DPs over time [Panel (C)]. Finally, to assess cross-study reproducibility of DPs, comparison of different sets of DPs derived from comparable dietary sources (at similar time points) is possible across centers from the same study, or across different studies representing potentially different populations or countries [Panel (D)]. In any of these 4 settings, confirming EFA-based DPs is possible through CFA, which assesses construct validity of DPs; results from the two approaches can be formally compared with suitable statistical tools [Panel (E)]. We re-classified the main findings from the articles included in the systematic review based on these definitions, no matter of the original definitions provided by the authors.

Supplemental Table 1. Definition of terms used in the current review and brief description of the statistical approaches used to assess these concepts in the current review[1]

Term	*General definition*	*Additional details within dietary pattern analysis*[2]	*Statistical method*[3]
Agreement	How close two measurements made on the same subject are? It is measured on the same scale as the measurements themselves. Agreement between measurements is a characteristic of the measurement methods involved (51)		Bland-Altman method with 95% LOA (limits are defined such that we expect that, in the long run, 95% of future differences between measurements made on the same subject will lie within the LOA) (15); Proportions of subjects classified into the same, adjacent, or opposite quantile category of score, or proportions of misclassified subjects (16); Kappa coefficient on score quantile categories (37); Sequence index plot (18)
Reliability	How inherent variability in the 'true' level of the quantity between subjects relates to the global variability of a phenomenon (variability in true levels plus variability in measurement error in observed measurements)? If reliability is high, measurement errors are small in comparison to the true differences between subjects, so that subjects can be relatively well distinguished (in terms of the quantity being measured) on the basis of the error-prone measurements. Conversely, if measurement errors tend to be large compared with the true differences between subjects, reliability will be low		Intraclass correlation coefficient between scores (52); Test-retest reliability on scores or on dominant food groups defining the identified dietary patterns (52) (see (5) for details)

	(51)		
Repeatability	How much is the variation in repeat measurements made on the same subject under identical conditions? Measurements are made by the same instrument or method, the same observer and they are made over a short period of time (over which underlying value considered constant). Variability in measurements made on the same subject in a repeatability study can then be ascribed only to errors due to the measurement process itself (51)		Pearson or Spearman or Kendall tau correlation coefficient between scores
Reproducibility	How much is the variation in measurements made on a subject under changing conditions? The changing conditions may be due to different measurement methods or instruments, measurements being made by different observers, or measurements being made over a longer period of time (within which the 'error-free' level of the variable could undergo non-negligible change) (51)	*Reproducibility across different statistical solutions:* Do different choices in the method used for the identification of DPs lead to similar sets of DPs? *Short-term reproducibility or reproducibility*: Are the sets of DPs derived at two administrations of the same dietary assessment tool to the same subjects within 1 year similar? Reproducibility of DPs is typically assessed following a previous assessment of reproducibility of a food-frequency questionnaire within a validation study *Long-term reproducibility or stability over time*: Are the sets of DPs derived at two or more administrations of the same dietary assessment tool to the same subjects over	Pearson (33) or Spearman (28) or Kendall tau correlation coefficient between scores; Intra-class correlation coefficient between scores (25); Congruence coefficient between loadings (11, 30)

		longer time periods (i.e., 2 years or more) similar? *Cross-study reproducibility:* Are the sets of DPs derived across centers (within the same study) or across different studies (potentially representing different populations or countries) similar?	
Validity	Does a test accurately measure what it claims to be measuring?		
Relative validity	Does a test compare well with a gold standard test? (53)	Are the sets of DPs derived on data from two different dietary sources similar? Relative validity of DPs is typically assessed following a previous assessment of relative validity of a food-frequency questionnaire against a gold standard tool within a validation study	Pearson (54) or Spearman (55) or Kendall tau (56) correlation coefficient between scores [crude or corrected (de-attenuated) for accounting for variation in time (57)]; Congruence coefficient between loadings (see (5) for details)
Construct validity	Does a test well measure the latent constructs that it is supposed to measure through operationalizations of the construct? (58)	Do the empirically derived DP scores resemble the latent DPs they should represent (in their composition and correlation with the other DPs)?	CFA (9, 27)

[1]ABBREVIATIONS: CFA: confirmatory factor analysis; DP: dietary pattern; LOA: limits of agreement

[2]See Supplemental Figure 1 for additional details

[3]For each statistical method mentioned, we provided an example study reference from the current review or from the companion one (5) to facilitate the association between research question and statistical method used to accomplish the objective

Supplemental Table 2. Cross-study reproducibility of *a posteriori* dietary patterns: details on dietary pattern composition[1]

Reference	*Location and Study*	*Dietary pattern composition*
Balder, 2003 (6)	Netherlands, Sweden, Finland, and Italy DIETSCAN (NLCS, SMC, ATBC, ORDET)	*From PCFAs based on unadjusted variables for energy intake*: (SALAD) VEGETABLE (common to all studies and different genders): high in raw leaf vegetables, tomatoes, carrots, cabbages and sometimes oil, poultry, rice, pasta and fish; PORK, PROCESSED MEAT, POTATOES (common to all studies and different genders): high in pork, processed meat, and potatoes; COOKED VEGETABLES (common to NCLS Ms and ORDET): high in cooked leaf vegetables, cabbages, legumes, and carrots; ALCOHOL (common to ATBC, SMC and ORDET): high in wine, beer, and spirits; SWEET AND/OR SAVORY SNACKS (common to NCLS Ms and Fs): high in savory snacks, nuts, sweets/candies, cakes/cookies; BROWN/WHITE BREAD SUBSTITUTION (common to NCLS Ms and Fs): high in bread substituters; *plus other 2 population-specific DPs not described in detail*
Castello, 2016 (12)	Spain EpiGEICAM, DDM-Spain	*From PCA on EpiGEICAM study data*: WESTERN: high in high-fat dairy products, processed meat, refined grains, sweets, energetic drinks and other convenience foods and sauces and low in low-fat dairy products and whole grains; PRUDENT: high in low-fat dairy products, vegetables, fruits, whole grains, and juices; MEDITERRANEAN: high in fish, vegetables, legumes, boiled potatoes, fruits, olives, and vegetable oil and low in juices *From PCA on DDM-Spain study data*: WESTERN: in addition with previous foods, low in white fish; PRUDENT: high in whole grains and juices but not on low-fat dairy products, vegetables and fruits; MEDITERRANEAN: high in some vegetables, legumes, potatoes, nuts, low-fat dairy products, sweets, and sugary and convenience foods, but not in fish, olive oil and fruits
Castello, 2016 (11)	Spain EpiGEICAM	*Dietary patterns based on original PCA from Castello and on reconstruction of loadings from Bessaoud, Adebamowo, and Terry*: WESTERN (Castello, Bessaoud, Adebamowo, Terry): high in high-fat dairy products (only cheese in Bessaoud), red and processed meat, refined grains, sweets, caloric drinks (not present in Bessaoud), and convenience food and sauces; Castello-PRUDENT, Bessaoud-MEDITERRANEAN, Adebamowo-PRUDENT, and Terry-HEALTHY: high in fish, fruits, and vegetables, also high in low-fat products ("dairy products" in Bessaoud); Castello-MEDITERRANEAN, Bessaoud-MEDITERRANEAN, Adebamowo-PRUDENT, and Terry-HEALTHY: high in fish, fruits, and vegetables, also high legumes (not present in Terry), nuts (not present in Bessaoud and Terry), and olive oil (not present in Adebamowo and Terry)

De Vito, 2019 (32)	USA, Italy, and Switzerland INHANCE	*From multi-study factor analysis on all the 7 available studies*: ANIMAL PRODUCTS AND CEREALS: high in total protein, zinc, phosphorus, riboflavin, sodium, niacin, thiamin, cholesterol, calcium, vitamin B6, iron, potassium, and total carbohydrates; ANTIOXIDANT VITAMINS AND FIBER: high in vitamin C, total fiber, total folate, potassium, total carotene, and vitamin B6; FATS: high in monounsaturated and polyunsaturated fatty acids, vitamin E, and saturated fatty acids *Study-specific DPs for the US studies only: 4 DPs with some variation but basically summarized as*: DAIRY PRODUCTS AND BREAKFAST CEREALS: high in calcium and low in niacin (or viceversa)
Judd, 2014 (26)	USA REGARDS	*From final PCA solution on the whole sample*: CONVENIENCE: high in mixed dishes with meat, pasta dishes, Mexican dishes, pizza, red meat, soup, fried potatoes, and Chinese dishes; PLANT-BASED: high in cruciferous, green leafy, dark yellow, and other vegetables, fruits, beans, and fish; SWEETS/FATS: miscellaneous sugar, desserts, bread, sweet breakfast foods, chocolate, candy, solid fats, and oils; SOUTHERN: high in added fats, eggs, fried food, organ meats, processed meats, and sugar-sweetened beverages; ALCOHOL/SALADS: high in salad dressing, green leafy vegetables, tomatoes, wine, butter, and liquor
Mannisto, 2005 (7)	Netherlands, Sweden, and Italy DIETSCAN (NLCS, SMC, ATBC, ORDET)	*From PCFAs on each study: common DPs*: VEGETABLES - VEG: high in vegetables, legumes, fruit, pasta, fish and oil; PORK, PROCESSED MEAT, POTATOES - PPP: high in pork, beaf, processed meats, potatoes, rice, poultry, liver, butter/low-fat margarine, pasta, and coffee; *plus other population-specific DPs not described in detail*
Moskal, 2014 (8)	Europe EPIC	*From overall PCA*: PC1: high in dietary fibre, vitamin C, beta-carotene and folate, low in saturated fatty acids, cholesterol, vitamin B12, retinol, and vitamin D; PC2: high in riboflavin, B6, folate, vitamin B12, vitamin C, beta-carotene, retinol, phosphorus, potassium and magnesium, low in starch; PC3: high in vitamin D, polyunsaturated fatty acids, thiamin, vitamin B6, and fibre, low in saturated fatty acids and retinol; PC4: high on calcium, total proteins, riboflavin, and phosphorus, low in polyunsaturated fatty acids and vitamin E
Schwerin, 1981 (43)	USA Ten-State Nutrition Survey (Ten-State), HANES I	*From PCA on Ten-State*: I: high in dairy products, and soups, and low in foods primarily sugar; II: high in nonsugary beverages and condiments and low in dairy products; III: high in cereals and grains, legumes and nuts, and eggs; IV: high in fruits, vegetables and juices, desserts and meats; V: high in poultry; VI: high in mixed dishes - protein, and shellfish; VII: high in fish and fats

Schwerin, 1982 (44)	USA Ten-State Nutrition Survey (Ten-State), HANES I, NFCS	*From PCA on Ten-State*: 7 DPs I: high in dairy products, and soups; II: high in nonsugary beverages and condiments; III: high in cereals and grains, legumes and nuts, and eggs; IV: high in vegetables and fruit, meats, and desserts; V: high in poultry; VI: high in mixed protein dishes and shellfish; VII: high in fish and fats and oils *From PCA on HANES I*: 8 DPs, of which 7 DPs similar to the Ten-State ones and 1 extra DP with greater consumption of sugary food and beverages *From PCA on NFCS*: 6 DPs, of which 5 were either identical over the decade or combinations of previous DPs

[1]ABBREVIATIONS: ATBC: Alpha-Tocopherol Beta-Carotene Cancer Prevention Study; DDM-Spain: Determinantes de la Densidad Mamográfica en España; DIETSCAN: DIETary patternS and CANcer in four European countries project; DP: dietary pattern; EPIC: European Prospective Investigation into Cancer and Nutrition; EpiGEICAM: Grupo Español de investigación en Cáncer de Mama; F: female; HANES: Health and Nutrition Examination Survey; INHANCE: International Head and Neck Cancer Epidemiology consortium; M: male; NFCS: Nationwide Food Consumption Survey; NLCS: Netherlands Cohort Study on diet and cancer; ORDET: Ormoni e Dieta nella Eziologia dei Tumori in Italy; PC: principal component; PCA: principal component analysis; PCFA: principal component factor analysis; REGARDS: Reasons for Geographic and Racial Differences in Stroke; SMC: Swedish Mammography Cohort

Supplemental Table 3. Stability over time of a *posteriori* dietary patterns: details on dietary pattern composition[1]

Reference	Location and Study	Dietary pattern composition
Asghari, 2012 (25)	Iran TLGS	*From PCAs on different dietary sources and time-points*: IRANIAN TRADITIONAL (common to all 4 dietary data): high in vegetables, fruits, potatoes, dairy products, legumes and nuts, whole grains, tea and coffee, olives, eggs, red meat, and organ meat; WESTERN (common to all 4 dietary data): high in carbonated drinks, salty snacks and salty vegetables, sugars, sweets, desserts, vegetable oil, animal fats, fast foods, poultry, fish and other seafood and refined grains; COMBINED (FFQ3 data only): high in potatoes, tea and coffee, vegetable oils, eggs, legumes and nuts, sugar, whole grains and salty snacks
Borland, 2008 (28)	UK SWS	*From PCAs at baseline and at follow-up*: PRUDENT DIET: high in vegetables, fruit, wholemeal bread, rice/pasta, yogurt, breakfast cereals, low in white bread, roast potatoes/chips, red/processed meat, full-fat milk, full-fat spread, crisps, confectionery, sugar, tea/coffee and Yorkshire puddings/pancakes, tinned vegetables, cakes and biscuits, and soft drinks; HIGH-ENERGY DIET: high in puddings, cakes/biscuits, potatoes/chips, vegetables, fruit, red/processed meat, fish, eggs, oils and full-fat spreads
Chen, 2015 (29)	Canada CCS, FFQVP	*From EFA on CCS study*: MEAT: high in red meat, cured/processed red meat, cured/processed meat, and mixed dishes; PLANT-BASED DIET: high in fruit, cruciferous vegetables, other green vegetables, beans, peas, other vegetables, tomato sauce, total cereals and grains, and whole grains; FISH: high in fish, processed fish, berries, and other local fruits, and low in cheese *From EFA on FFQVP study*: as above for the MEAT and FISH DPs, but the PLANT-BASED DP becomes: VEGETABLES/FRUITS: high in greens, tomato sauce, berries, and other vegetables; *plus an additional DP*: GRAINS: high in whole grains, cereal, grains and low in beer, white wine, and coffee

Crozier, 2009 (15)	UK SWS	*From PCAs at 3 time-points*: PRUDENT: high in fruit and vegetables, whole-meal bread, rice and pasta, yogurt, and low in chips and roast potatoes, sugar, white bread, processed meat, full-fat dairy products, crisps, Yorkshire puddings and savory pancakes, confectionery, and tea and coffee; HIGH-ENERGY DIET: high in fruit and vegetables, puddings, meat and fish, eggs and egg dishes, cakes and biscuits, full-fat spread, potatoes, crisps, and confectionery; *plus extra DPs not shared across time-points and not described in detail (low total variance explained and less interpretable)*
Cucò, 2006 (30)	Spain NA	*From PCFAs at the 5 time-points with some variation*: SWEETENED BEVERAGES AND SUGARS: high in sweetened beverages and sugars, and low in fresh fruit, vegetables, roots and tubers (signs inverted in some of the time-points); VEGETABLES AND MEAT (not present in the postpartum period): high in vegetables, roots and tubers, red meat, cured cold meats, olive oil, and eggs
Cutler, 2009 (31)	USA (Minnesota) EAT	*From PCFA at Time 1*: across cohort and gender, same set of 4 DPs identified: VEGETABLE: high in zucchini, squash, eggplant, kale and greens, spinach, peas and lima beans; FRUIT: high in oranges and grapefruit, apples and apple sauce, pears, grapes, bananas, strawberries, cantaloupe and melons, peaches, and plums and apricots; SWEET/SALTY SNACK FOOD: high in chocolate bars, other candy bars, candy with chocolate, brownies, cake, potatoes chips, and nachos; STARCHY FOOD: high in English muffins/bagels, grilled cheese, pancakes, and crackers for 3 subgroups, and high in mashed potatoes, lasagna, pretzels, macaroni and cheese, and spaghetti with sauce for 2 subgroups; *From PCFA at Time 2*: previous 4 DPs, not identical anymore across cohort and gender (except for young girls), but fairly similar: VEGETABLE AND FRUIT DPs: combined in older boys and girls and separate in younger boys (and girls); SWEET/SALTY SNACK FOOD: identical across cohort and gender; STARCHY FOOD: in younger and older girls only; FAST FOOD: high in hamburgers, French fries, fried food, nondiet soda; identified in all age/sex groups except young girls
Dekker, 2013 (10)	Netherlands Doetinchem Cohort Study	*From CA on each of the 3 surveys*: HIGH-FIBER BREAD: high percentage of total energy from high-fibre bread, cakes and cookies, and cheese; LOW-FIBER BREAD: high percentage of total energy from low-fibre bread, sugar-sweetened beverages, other alcoholic drinks and fries

Fung, 2001 (33)	USA HPFS	*From PCFAs at the 3 time-points*: PRUDENT: high in fruit, vegetables, poultry, fish, whole grains, and legumes; WESTERN: high in red and processed meat, French fries, eggs, high-fat dairy products, sweets, and refined grains
Gerdes, 2002 (34)	Denmark MONICA	*From PCFAs at each of the 3 surveys*: COARSE BREAD: high in coarse bread; BAKED GOODS AND SWEETS: high in cakes and biscuits, jam, honey, candy, ice cream, and soda; FRUIT AND VEGETABLES: high in fruit, juice, vegetables, and cheese; MEAT, POTATOES AND FAT: high in meat, sausages, potatoes, butter, fat, and margarine; PASTA AND RICE: high in Mediterranean and Asian cooking; BREAKFAST: high in porridge, oatmeal, milk, yogurt, jam, and honey
Lopez-Garcia, 2004 (35)	USA NHS	*From PCFAs at the 2 time-points and on the average consumption from the 2 time-points*: PRUDENT: high in vegetables, fruit, legumes, whole grains, fish, and poultry; WESTERN: high in red meat, processed meat, refined grains, sweets, desserts, French fries, and high-fat dairy products
Malik, 2012 (36)	USA NHS II	*From EFAs at the 5 time-points with some variation in the 2003 DPs*: PRUDENT: high in vegetables, fruit, legumes, fish, and better-quality grains, low in snacks and soda; WESTERN: high in desserts, snacks, processed meat, French fries, and refined grains, and low in vegetables, fruit, and fish
Mikkila, 2005 (16)	Finland Cardiovascular Risk in Young Finns Study	*From PCFAs at the 3 time-points with some variation described*: PATTERN 1: high in rye, potatoes, milk, butter, sausages, coffee (at all time-points), low in fruit and berry, and other dairy products (in 1980 and 2001); PATTERN 2: high in rye, vegetables, legumes and nuts, tea, rye, cheese and other dairy products (at all time-points), and alcoholic beverages (in 2001); *plus one extra DP not described in detail but not easily interpretable*

Mishra, 2006 (37)	UK Medical Research Council National Survey of Health and Development (1946 British Birth Cohort)	*From EFAs on 1999 data, but similar to 1982 and 1989 data*: Among Fs; ETHNIC FOOD AND ALCOHOL: high in Indian and Chinese meals, rice, pasta, oily fish and shellfish, olive oil, some vegetables, and alcoholic beverages; MEAT, POTATOES AND SWEET FOODS: high in red meat, bacon, ham, potatoes, sweet pies, cakes, puddings and desserts, and low in pasta, and skimmed milk; FRUIT, VEGETABLES AND DAIRY: low-fat and reduced-fat dairy products, fruit, some vegetables and whole-meal bread, and low in meat, meat products, and white bread Among Ms; ETHNIC FOOD AND ALCOHOL: high in Indian and Chinese meals, rice, pasta, shellfish, olives, some vegetables and legumes, and alcoholic beverages, and low in meat pies, fried chips, and animal fats; MIXED: high in many fruits and vegetables, low-fat/low-calorie yogurt and soya milk and cakes, sweet biscuits, sweet pies, puddings, desserts, confectionery, and ice cream
Newby, 2006 (23)	Sweden SMC	*From PCFA at both time-points (1987 and 1997) and confirmed with CFA at both time-points (1987 and 1997)*: HEALTHY: high in vegetables, fruit, whole grains, fruit juice, and cereal; WESTERN/SWEDISH: high in meat, processed meat, liver, refined grains, and potatoes; ALCOHOL: high in wine, spirits, snacks beer, and chocolate; SWEETS: high in sweet baked goods, chocolate, sugary foods, dairy desserts, soda, fruit soup, and refined grains; *plus 2 extra DPs not shared among the 2 time-points*
Newby, 2006 (22)	Sweden SMC	*From PCFA at both time-points (1987 and 1997) and confirmed with CFA at both time-points (1987 and 1997)*: with some variation in the Healthy DP (seafood, poultry, and eggs also contributed to HEALTHY DP in 1987, whereas legumes and soy products contributed to HEALTHY DP in 1997) HEALTHY: high in vegetables, fruit, whole grains, fruit juice, and cereal; WESTERN/SWEDISH: high in meat, processed meat, liver, refined grains, and potatoes; ALCOHOL: high in wine, spirits, snacks beer, and chocolate; SWEETS: high in sweet baked goods, chocolate, sugary foods, dairy desserts, soda, fruit soup, and refined grains; *plus 2 extra DPs not shared among the 2 time-points*
Nimptsch, 2014 (38)	USA NHS II	*From EFAs at the 5 time-points*: PRUDENT: high in vegetables, fruit, better-quality grains, fish, and poultry; WESTERN: high in desserts and sweets, snack foods, red and processed meat, French fries, and refined grains

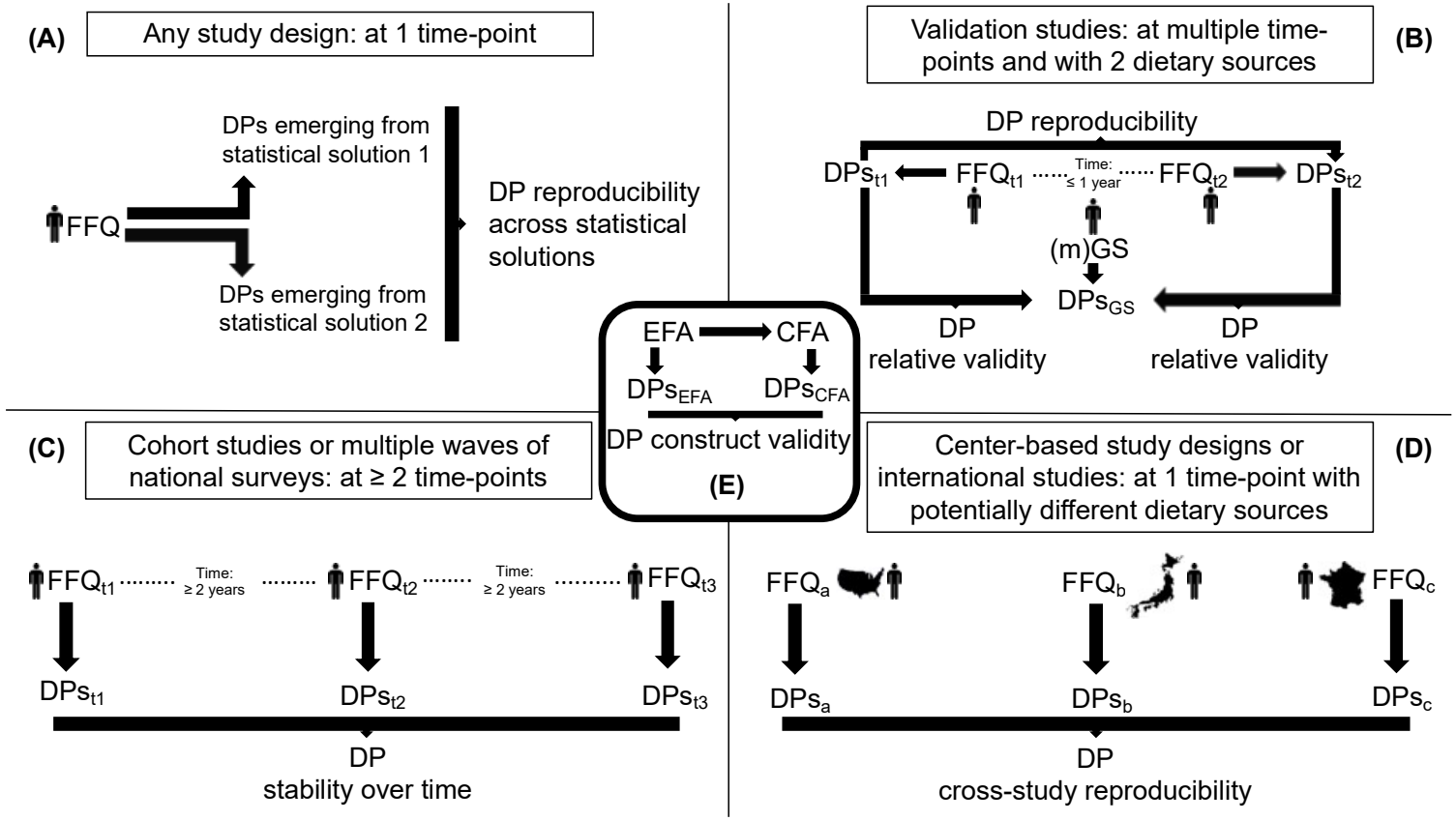
Northstone, 2005 (39)	UK ALSPAC	*From PCAs at 2 time points*: JUNK: high in high-fat and sugar content, processed and convenience foods; TRADITIONAL: high in meat, potatoes and vegetables; HEALTH-CONSCIOUS: high in vegetarian style foods, rice, pasta, cheese, salad, fish, and fruit
Northstone, 2013 (18)	UK ALSPAC	*From CAs at all time-points, in order of size*: PROCESSED: higher mean consumption of processed meat, pies and pasties, coated and fried chicken and white fish, pizza, chips, baked beans and tinned pasta, chocolate, sweets, sugar and diet and regular fizzy drinks; HEALTHY: higher mean consumption of non-white bread, reduced fat milk, cheese, yogurt and fromage frais, butter, breakfast cereal, rice, pasta, eggs, fish, vegetable and vegetarian dishes, soup, salad, legumes, fruit, crackers and crispbreads, high-energy-density sauces (e.g. mayonnaise), fruit juice, and water; TRADITIONAL: higher mean consumption of red meat, poultry, potatoes, vegetables, starch-based products (e.g. Yorkshire pudding), low-energy-density sauces (e.g. gravy), puddings, tea and coffee; PACKED LUNCH: higher mean consumption of white bread, margarine, ham and bacon, sweet spreads (e.g. honey), salty flavourings (e.g. yeast extract), crisps, biscuits, diet squash, tea and coffee
Northstone, 2008 (17)	UK ALSPAC	*From PCA on 4 time-points*: PROCESSED (at all time-points): high in high-fat and sugar content foods and processed and convenience foods; TRADITIONAL (at all time-points): high in meat, poultry, potatoes and vegetables; HEALTH CONSCIOUS (at 3, 4, and 7 ys only): high in salads, fruit, vegetables, fish, pasta and rice; HEALTH CONSCIOUS/VEGETARIAN (at 9 ys only): high in salads, fruit, vegetables, fish, pasta and rice, but also high in meat substitutes, pulses, nuts and vegetarian pies; SNACK (at 3 ys only): high in cheese, fruit, puddings, cakes, biscuits, and crisps
Northstone, 2008 (40)	UK ALSPAC	*From PCA on pregnancy data*: HEALTH-CONSCIOUS: high in salad, fresh fruit, rice, pasta, fish, pulses, and non-white bread; TRADITIONAL (British): high in all types of vegetables, some red meat, and poultry; PROCESSED: high in meat pies, sausage and burgers, fried foods, pizza, and chips; CONFECTIONARY: high in chocolate, sweets, biscuits, cakes and other puddings; VEGETARIAN: high in meat substitutes, pulses, nuts, and herbal tea; *From PCA on 4-y follow-up data*: TRADITIONAL DP lost, HEALTH-CONSCIOUS similar, the other 3 DP virtually identical in the dominant FG across time

Prevost, 1997 (41)	UK HALS	*From PCAs at HALS1 and HALS2 and also similar for Ms and Fs*: COMPONENT 1 (HIGH IN FRUIT AND VEGETABLES, LOW IN FAT): high in fresh fruit, salads, brown bread, fruit juice, green vegetables, spread (low-fat), milk (semi-skimmed), other vegetables, and root vegetables, low in chips, fried foods, and processed meats; COMPONENT 2 (HIGH IN ENERGY-DENSE FOODS): high in puddings/pies, cake, potatoes, biscuits, preserves, pulses, carcass meat, root vegetables, cream, cooked fruit, confectionery, green vegetables, milk, eggs, light desserts; COMPONENT 3 (HIGH IN CONVENIENCE FOODS): high in crisps, soft drinks, chips, fried food, coffee, pasta/rice, processed meat, low in tea and preserves; COMPONENT 4 (HIGH IN SUGARY FOODS, LOW IN VEGETABLES): high in confectionery, biscuits, cake, and low in green vegetables, root vegetables, pulses, other vegetables, and potatoes
Schulze, 2006 (42)	USA NHS II	*From EFAs at all time-points*: PRUDENT: high in fruits, vegetables, whole grains, fish, poultry, and salad dressings; WESTERN: high in red and processed meat, refined grains, sweets and desserts, and potatoes
Togo, 2004 (27)	Denmark MONICA	*From CFA among Ms, at both baseline and follow-up*: GREEN: high in wheat bread and rye bread with whole grains and/or bran; raw and boiled vegetables, fruit, rice, cheese, fish, milk products and low in white (wheat) bread; SWEET: high in cake, biscuits, or other baked goods, candy or chocolate, soft drink or ice-cream, and jam/marmalade or honey; TRADITIONAL: high in meat, paté and meat for bread, potatoes, white (wheat) bread, sausage, butter, lard and hard margarine, and eggs; *From CFA among Fs, at both baseline and follow-up*: GREEN: same as for Ms; SWEET-TRADITIONAL: high in candy or chocolate, cake, biscuits, or other baked goods, paté and meat for bread, white (wheat) bread, butter, lard and hard margarine, soft drink or ice-cream, jam/marmalade or honey, potatoes, meat, and sausage
van Dam, 2002 (45)	USA HPFS	*From PCFAs at the 2 time-points*: PRUDENT: high in vegetables, legumes, fruit, whole grains, fish, and poultry; WESTERN: high in red meat, processed meat, refined grains, French fries, high-fat dairy products, sweets and desserts, high-sugar drinks, and eggs

Weismayer, 2006 (9)	Sweden SMC	*From EFAs at baseline and follow-up and confirmed by CFAs at baseline and follow-up*: HEALTHY: high in fruits, tomatoes, vegetables, cereal, and fish; WESTERN: high in meat, processed meat, fried potatoes, soft drinks, and sweets; ALCOHOL: high in beer, wine, and liquor consumption as well as snack consumption; *plus extra DPs difficult to interpret or dominated by only 1 high loading*

[1]ABBREVIATIONS: ALSPAC: Avon Longitudinal Study of Parents and Children; CA: cluster analysis; CCS: Case-Control Study, here intended as the full name of one of the included studies and not as the case-control study design; CFA: confirmatory factor analysis; DP: dietary pattern; EAT: Eating Among Teens; EFA: exploratory factor analysis; F: female; FFQ: food-frequency questionnaire; FFQ1/FFQ2/FFQ3: food-frequency questionnaire at time 1/2/3; FFQVP: Food-Frequency Questionnaire Validation Project; FG: food groups; HALS: Health and Lifestyle Survey; HPFS: Health Professionals Follow-up Study; M: male; MONICA: MONItoring of trends and determinants in CArdiovascular Disease; NA: not available; NHS: Nurses' Health Study; PCA: principal component analysis; PCFA: principal component factor analysis; SMC: Swedish Mammography Cohort; SWS: Southampton Women's Survey; TLGS: Teheran Lipid and Glucose Study; y: year

Supplemental Figure 1. Schemes of dietary pattern identification processes related to the assessment of their reproducibility and validity. Specifically, reproducibility and/or validity of dietary patterns can be assessed in the following set-ups: Panel (A): at one time point and with one dietary source; Panel (B): at multiple time points and with two dietary source, Panel (C): at multiple time points; Panel (D): across centers from the study or across different studies. All of these settings may include confirmation of the identified dietary patterns with confirmatory factor analysis [Panel (E)][1]



[1]ABBREVIATIONS: CFA: confirmatory factor analysis; DPs: dietary patterns; EFA: exploratory factor analysis: FFQ: food-frequency questionnaire; GS: gold standard dietary assessment tool; mGS: mean of intakes from multiple administrations of the same gold standard tool

Supplemental References
[In addition to references cited in the main text]

51. Bartlett JW, Frost C. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. Ultrasound Obstet Gynecol 2008;31(4):466–75.
52. Ryman TK, Boyer BB, Hopkins S, Philip J, O'Brien D, Thummel K, Austin MA. Characterising the reproducibility and reliability of dietary patterns among Yup'ik Alaska Native people. Br J Nutr 2015;113(4):634-43. doi: 10.1017/S0007114514003596.
53. Willett W. Nutritional Epidemiology. Oxford University Press; 2012, Oxford, UK.
54. Beck KL, Kruger R, Conlon CA, Heath AL, Coad J, Matthys C, Jones B, Stonehouse W. The relative validity and reproducibility of an iron food frequency questionnaire for identifying iron-related dietary patterns in young women. J Acad Nutr Diet 2012;112(8):1177-87. doi: 10.1016/j.jand.2012.05.012.
55. Ambrosini GL, O'Sullivan TA, de Klerk NH, Mori TA, Beilin LJ, Oddy WH. Relative validity of adolescent dietary patterns: a comparison of a FFQ and 3 d food record. Br J Nutr 2011;105(4):625-33. doi: 10.1017/S0007114510004137.
56. Bountziouka V, Tzavelas G, Polychronopoulos E, Constantinidis TC, Panagiotakos DB. Validity of dietary patterns derived in nutrition surveys using a priori and a posteriori multivariate statistical methods. Int J Food Sci Nutr 2011;62(6):617-27. doi: 10.3109/09637486.2011.561783.
57. Hu FB, Rimm E, Smith-Warner SA, Feskanich D, Stampfer MJ, Ascherio A, Sampson L, Willett WC. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. Am J Clin Nutr 1999;69(2):243-9.
58. Strauss ME. Introduction to the special section on construct validity of psychological tests: 50 years after Cronbach and Meehl (1955). Psychol Assess 2005;17(4):395. doi: 10.1037/1040-3590.17.4.395.