



National Institute for Public Health  
and the Environment  
*Ministry of Health, Welfare and Sport*

# Dutch tool(s) for food matching

Richfields Workshop on Dietary  
Assessment and Food Matching  
Tools, April 8<sup>th</sup> 2016

Ido Toxopeus  
Marga Ocké  
Susanne Westenbrink

Short intro: Dutch tool for food matching  
April 8th 2016



## Dutch Food Consumption Survey (DFCS)

- General population (aged 3-80)
- Approximately 4000 participants (1000 / year)
- 2 x 24hrs-recall
- Globodiet
- Current data collection: 2012 – 2017
  
- Previously
  - 2007-2010 General population (aged 7-70)
  - 2010-2011 Elderly population (aged 70+)
  
- Consumed items are unique in  
Globodiet FoodNum – Foodgroup - Name – Brand Name – Facets




## Dutch Food Composition Database (NEVO)

- Dutch Food Composition Database used by researchers and dieticians.
- Approximately 2500, mostly generic, food items

energie en macronutriënten (per 100g voedingsmiddel)

eiwitten koolhydr vetten verz vetz eo vetz mo v

	NEVO-code	Voedingsmiddel	Food item	Energie (kJ)	Energie (kcal)	Eiwit totaal (g)	Koolhydraten
	455	Pindakaas	Peanut butter	2792	675	22.4	13.6



## Linking DFCS - NEVO

- DFCS 2007-2010 – NEVO: manually using Excel files
- DFCS 2010-2011 – NEVO: manually using Excel files
  
- -> 30782 unique consumed items linked to NEVO 1639 unique generic food items
  
- First 2 years of data collection from current DFCS: 21851 unique consumed items already!



## Linking DFCS - NEVO

- Food Matching Tool helps to use history
  - Selects consumed items from historic data with exact match on Globodiet FoodNum – Food group - Name – Brand Name – Facets  
If found: use the same NEVO code
  - Saves time & increases efficiency and consistency
    - › Allows for filtering and sorting of items to match
    - › Match multiple consumed items in single action
    - › User can (de)select fields used for matching
    - › Allows for documentation, follow-up actions
    - › Review of decisions



# DEMO – select food group to work on

SelecteerVCPFoodgroup

Kies een beschikbare VCP Foodgroup om te bewerken

vcp_naam	vcp_foodgroup	Aant gekoppelde	Aant te koppelen
VCPBasis_12-17	05 - Dairy products and substitutes	2082	2294

waarvan

2082 alleen aan Nevo

0 aan NEVO en FoodEx2

Databeheer

Kies deze foodgroup en ga naar formulier om koppeling te beginnen

Kies deze foodgroup en bekijk producten alleen aan NEVO gekoppeld

Kies deze foodgroup en bekijk producten aan NEVO én FoodEx2 gekoppeld

Sluit applicatie af





# DEMO – select food item to match ... no matches

KoppelVCPFoodItem\_NEVOProd

Kies een te koppelen VCP fooditem uit lijst met nog niet gekoppelde fooditems

vcp	foodgr	foodn	name	brandname	facets_aggr	aant
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	alpro	0701,1102,1201	1
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	alpro amandel	0701,1199,1201	1
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	alpro zonder s	0701,1101,1102,1201	1
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	provamel	0701,1199,1201	2
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	provamel	0701,1100,1201	1
VCPBasis_1	0506	02675	Creme brulee		0700,0801,10AB,1200	1
VCPBasis_1	0506	02675	Creme brulee	(restaurant)	0702,0800,10AB,1201	1
VCPBasis_1	0506	02675	Creme brulee	albert heijn	0702,0801,10AB,1201	1

te koppelen VCP fooditem

foodgroup	foodnum
<input type="text" value="0506"/>	<input type="text" value="02675"/>
name	
<input type="text" value="Creme brulee"/>	
brandname	
<input type="text"/>	
facetstring_aggr	
<input type="text" value="0700,0801,10AB,1200"/>	
aantkeergegeten	<input type="text" value="1"/>

Zoek eerdere koppelingen

Kies een NEVO Product om mee te koppelen op basis van de suggesties van eerdere koppelingen:

Pas aan

Zoek naar exacte overeenkomst op:  Selectie vast  foodgroup  foodnum  name  brandname  facetten:  Allen/Geen  01  02  03  04  05  06  07  08  09

vcp	foodgr	foodn	name	brandname	facets	aant	koppelstat	nevonaam	nevoprodgr	nevoo



# DEMO – ignore facets ... find exact matches

KoppelVCPFoodItem\_NEVOProd

Kies een te koppelen VCP fooditem uit lijst met nog niet gekoppelde fooditems

vcp	foodgr	foodn	name	brandname	facets_aggr	aant
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	alpro	0701,1102,1201	1
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	alpro amandel	0701,1199,1201	1
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	alpro zonder s	0701,1101,1102,1201	1
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	provamel	0701,1199,1201	2
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	provamel	0701,1100,1201	1
VCPBasis_1	0506	02675	Creme brulee		0700,0801,10AB,1200	1
VCPBasis_1	0506	02675	Creme brulee	(restaurant)	0702,0800,10AB,1201	1
VCPBasis_1	0506	02675	Creme brulee	albert heijn	0702,0801,10AB,1201	1

te koppelen VCP fooditem

foodgroup  foodnum

name

brandname

facetstring\_aggr

aantkeergegeeten

Zoek eerdere koppelingen

Kies een NEVO Product om mee te koppelen op basis van de suggesties van eerdere koppelingen:

Zoek naar exacte overeenkomst op:  Selectie vast  foodgroup  foodnum  name  brandname  facetten:  Allen/Geen  01  02  03  04  05  06  07  08  09

vcp	foodgr	foodn	name	brandname	facets	aant	koppelstat	nevonaam	nevoprodr	nevoc
VCPBasis_12-17	0506	02675	Creme brulee		0702,0800,10AB,1200	2	Voorlopig	Nevo_13	Melk en melk	933
VCPOuderen_10-1	0506	02675	Creme brulee		0604,0605,0621,0700,0800,10AB	1	Voorlopig	Nevo_13	Melk en melk	933
VCPBasis_07-10	0506	02675	Creme brulee		0604,0702,0801,10AA	1	Voorlopig	Nevo_11	Melk en melk	933
VCPBasis_07-10	0506	02675	Creme brulee		0600,0604,0702,0800,1000	1	Voorlopig	Nevo_11	Melk en melk	933
VCPBasis_07-10	0506	02675	Creme brulee		0600,0702,0800,1000	1	Voorlopig	Nevo_11	Melk en melk	933
VCPBasis_07-10	0506	02675	Creme brulee		0603,0702,0800,10AB,1200	1	Voorlopig	Nevo_11	Melk en melk	933

Pas aan





# DEMO – select food item/NEVO code (all the same)

KoppelVCPFoodItem\_NEVOProd

Kies een te koppelen VCP fooditem uit lijst met nog niet gekoppelde fooditems

vcp	foodgr	foodn	name	brandname	facets_aggr	aant
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	alpro	0701,1102,1201	1
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	alpro amandel	0701,1199,1201	1
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	alpro zonder s	0701,1101,1102,1201	1
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	provamel	0701,1199,1201	2
VCPBasis_1	0502	03364	Amandeldrank Alpro en soortgeli	provamel	0701,1100,1201	1
VCPBasis_1	0506	02675	Creme brulee		0700,0801,10AB,1200	1
VCPBasis_1	0506	02675	Creme brulee	(restaurant)	0702,0800,10AB,1201	1
VCPBasis_1	0506	02675	Creme brulee	albert heijn	0702,0801,10AB,1201	1

te koppelen VCP fooditem

foodgroup:  foodnum:

name:

brandname:

facetstring\_aggr:

aantkeergegeten:

Verrijking(en)

Indicator Rauw

Opmerking bij koppeling:

Kies Status koppeling:

Volgende stap: selecteer het volgende VCP product

Kies een NEVO Product om mee te koppelen op basis van de suggesties van eerdere koppelingen:

Zoek naar exacte overeenkomst op:  Ververs  Selectie vast  foodgroup  foodnum  name  brandname  facetten:  Allen  Geen  01  02  03  04  05  06  07  08  09  10  11  12  13  14  15  16  17  18  19

vcp	foodgr	foodn	name	brandname	facets	aant	koppelstat	nevo naam	nevo prodgr	nevo code	productnaam	indrauv	opmer
VCPBasis_12-17	0506	02675	Creme brulee		0702,0800,10AB,1200	2	Voorlopig	Nevo_13	Melk en melk	933	Pudding caramel-		als pudding caram
VCPouderen_10-1	0506	02675	Creme brulee		0604,0605,0621,0700,0800,10AB	1	Voorlopig	Nevo_13	Melk en melk	933	Pudding caramel-		vr, van, room in cc
VCPBasis_07-10	0506	02675	Creme brulee		0604,0702,0801,10AA	1	Voorlopig	Nevo_11	Melk en melk	933	Pudding caramel-		zie boven
VCPBasis_07-10	0506	02675	Creme brulee		0600,0604,0702,0800,1000	1	Voorlopig	Nevo_11	Melk en melk	933	Pudding caramel-		als pudding caram
VCPBasis_07-10	0506	02675	Creme brulee		0600,0702,0800,1000	1	Voorlopig	Nevo_11	Melk en melk	933	Pudding caramel-		zie boven
VCPBasis_07-10	0506	02675	Creme brulee		0603,0702,0800,10AB,1200	1	Voorlopig	Nevo_11	Melk en melk	933	Pudding caramel-		als pudding caram



# DEMO – Inspect details, confirm or change

Doorzoek\_Gekoppeld

Doorzoek producten die alleen NEVO zijn gekoppeld

vcp	foodgr	foodnu	name	brandname	facets_aggr	facets_aggr_corr	orig_facets
VCPBasis_12-17	0506	02675	Creme brulee		0700,0801,10AB,1200	0700,0801,10AB,1200	
VCPBasis_12-17	0506	02675	Creme brulee		0702,0800,10AB,1200	0702,0800,10AB,1200	
VCPBasis_12-17	050701	00442	Creme fraiche		0201,0400,0700,0800,1200	0201,0400,0700,0800,1200	
VCPBasis_12-17	050701	00442	Creme fraiche		0201,0400,0701,0800,1200	0201,0400,0701,0800,1200	
VCPBasis_12-17	050701	00442	Creme fraiche		0201,0400,0701,0801,1200	0201,0400,0701,0801,1200	
VCPBasis_12-17	050701	00442	Creme fraiche		0201,0408,0700,0800,1200	0201,0408,0700,0800,1200	
VCPBasis_12-17	050701	00442	Creme fraiche		0201,0408,0701,0801,1200	0201,0408,0701,0801,1200	
VCPBasis_12-17	050701	00442	Creme fraiche	(aldi)	0201,0413,0701,0801,1201	0201,0413,0701,0801,1201	
VCPBasis_12-17	050701	00442	Creme fraiche	albert heijn	0201,0400,0700,0801,1201	0201,0400,0700,0801,1201	
VCPBasis_12-17	050701	00442	Creme fraiche	albert heijn	0201,0408,0701,0800,1201	0201,0408,0701,0800,1201	
VCPBasis_12-17	050701	00442	Creme fraiche	albert heijn	0201,0408,0701,0801,1201	0201,0408,0701,0801,1201	
VCPBasis_12-17	050701	00442	Creme fraiche	c1000	0201,0408,0701,0801,1201	0201,0408,0701,0801,1201	
VCPBasis_12-17	050701	00442	Creme fraiche	campina	0201,0408,0701,0801,1201	0201,0408,0701,0801,1201	
VCPBasis_12-17	050701	00442	Creme fraiche	de zaanse hoe	0201,0408,0700,0800,1201	0201,0408,0700,0800,1201	
VCPBasis_12-17	050701	00442	Creme fraiche	de zaanse hoe	0201,0408,0701,0801,1201	0201,0408,0701,0801,1201	
VCPBasis_12-17	050701	00442	Creme fraiche	jumbo	0201,0408,0701,0801,1201	0201,0408,0701,0801,1201	
VCPBasis_12-17	050701	00442	Creme fraiche	melkan	0201,0408,0701,0800,1201	0201,0408,0701,0800,1201	

Record: 1 of 2083 No Filter Search

VCP product

02675 Creme brulee  
0700,0801,10AB,1200  
 Rauw

als pudding caramel, net als bij eerdere VCPs, indien veel gebruikt evt nieuwe code

Ontkoppel

Bekijk (en wijzig) details

Ga terug

NEVO koppeling

933  
Pudding caramel-  
Voorlopig

nevo\_controle

FoodEx2 koppeling

FE2 controle

Sla controle aanvinkvakjes op voor alle geselecteerde regels (leeg = nee)

Ga naar koppeling FoodEx2

controle\_memo

Bewerk

Sla op of Annuleer

TOXOPEUI 20160405: nevo 933 Voorlopig.




## Work in progress

- Using Food Matching Tool
- Preliminary exploration of statistical matching technique
  - This technique also tries to find the consumed item in historic data that matches the current consumed item best.
    - › Look for historic consumed items that have an exact match on as many fields as possible with current consumed item
    - › Look for largest overlap on fields that do not show exact match
    - › Find historic consumed item with best 'evidence'.
    - › Use the NEVO code that was connected to this item.



## Work in progress: statistics to match

- Use historic matches
  - Fields available for matching:
    - Group-top
    - Group-sub1
    - Group-sub2
    - Brandname-NL
    - Name-EN
    - Name-NL
    - Facets
- good predictor
- 
- poor predictor



## Work in progress: statistics to match

- Select a subgroup of items that show an exact match on a maximum number of fields with the new (to match) item.
- Eg. Find 170 items in historic matches with an exact match on 3 fields, e.g.:
  - Group-top
  - Group-sub1
  - Group-sub2
  - Brandname-NL
  - Name-EN
  - Name-NL
  - Facets





## Work in progress: statistics to match

- Subgroup example
  - Item to match

Group-Top	Group-Sub1	Group-Sub2	Name-NL	Name-EN	Brandname-NL	Facets
5	8	ND	Koffiecreamer	Creamer	Deen	1201

### – Subgroup

Group-Top	Group-Sub1	Group-Sub2	Name-NL	Name-EN	Brandname-NL	Facets
5	8	ND	koffiemelk	Milk for coffee	Jumbo	802, 1201
5	8	ND	koffiemelk	Milk for coffee	Autobar	800, 1201
5	8	ND	Koffiemelk met plantaardig vet	Thickened milk for coffee vegetal fat		801,1201
5	8	ND	Koffiemelk/room/creamer poeder n.s.	Thickened milk for coffee, powdered n.s.	Landhof	800,1201



## Work in progress: statistics to match

- Look for the same particles in fields that do NOT match exactly
  - Split contents of fields into particles:
    - > Name-EN: Milk for coffee
    - > Facets: 800, 1201
  - Calculate overlap between non-matching fields (matching particles) :  $\frac{1}{2} \left( \frac{|A \cap B|}{|A|} + \frac{|A \cap B|}{|B|} \right)$

	<b>Product X</b>	<b>Product Y</b>	<b>overlap</b>
Name-EN	Milk for coffee	Coffee creamer	0.42
Facets	800, 1201	800, 1200	0.5



## Work in progress: statistics to match

- Do this for all items in the subgroup
- Use logistical regression technique to calculate an evidence weight



## Work in progress: statistics to match

- First preliminary results

<b>predicted NEVO code</b>	<b>evidence</b>	<b>type of match</b>	<b>size of subgroup</b>	<b>TRUE NEVO code</b>	<b>correct prediction or not</b>	<b>proportion of correct predictions</b>
1150	0.999225	logistic.regression	174	1150	1	0.726506323
1150	0.999159	logistic.regression	174	1150	1	
1150	0.999087	logistic.regression	174	1150	1	
1150	NA	closest.possible	1	1150	1	
1150	NA	closest.possible	1	1150	1	
982	0.946883	logistic.regression	293	982	1	
2323	0.613727	logistic.regression	43	2323	1	
2621	0.437664	logistic.regression	43	121	0	
2324	0.613727	logistic.regression	43	2324	1	
2324	0.613727	logistic.regression	43	982	0	
2324	0.613727	logistic.regression	43	121	0	
2324	0.613727	logistic.regression	43	2576	0	



## What's next?

- Further develop optimization techniques for matching:
  - Fine tune statistical method:
    - › what is best way to split fields into particles?
    - › How to deal with negations: “no fat”
    - › How to deal with ‘knowledge rules’: e.g.: facet-descriptors irrelevant for nutrient composition: Fried=stir fried=shallow fried
  - Also provide alternatives with lower ‘evidence’ for expert judgement.





## What's next?

- Further develop optimization techniques for matching:
  - Use both:
    - › historical matches (when available) based on food consumption field matching
    - › Food characteristics from food composition database, e.g.
      - Food group
      - Name
      - Descriptors, such as LanguaL, FoodEx2
  - Use barcodes (e.g. GS1) in consumption survey requires matching food composition to brand products database
    - Use also nutrient composition



## Food for thought

- In consumption survey can we link with food composition earlier in the process, during the interview?