



CONSTRUCTING A DATA WAREHOUSE FOR NUTRITION CLINICAL STUDIES

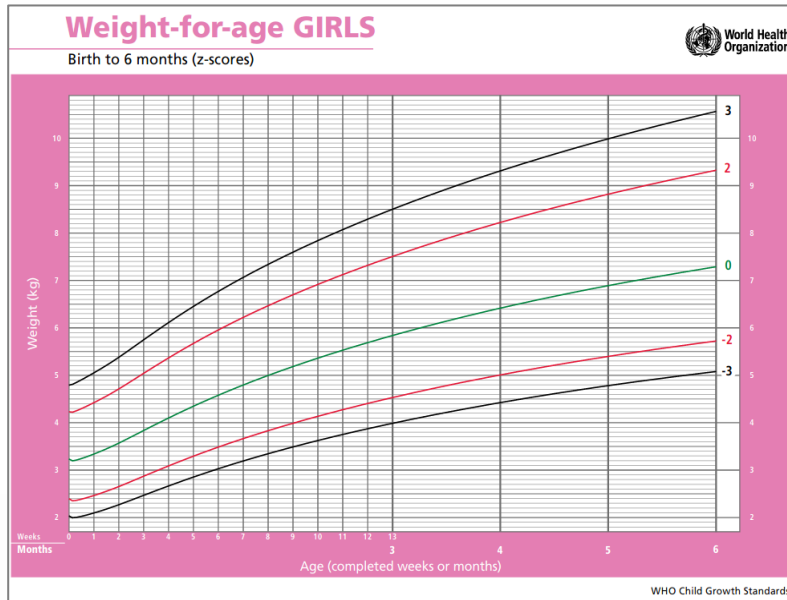
Paul Vervuren

PhUSE Barcelona
October 10th, 2016

**NUTRICIA
RESEARCH**



CLINICAL DATA WAREHOUSE PROJECT



Learn and Innovate

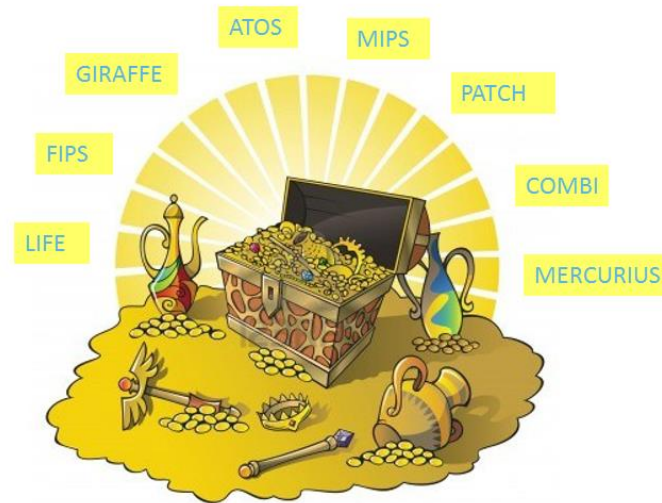
- Design new studies

- Reference data

DNR Growth curve?

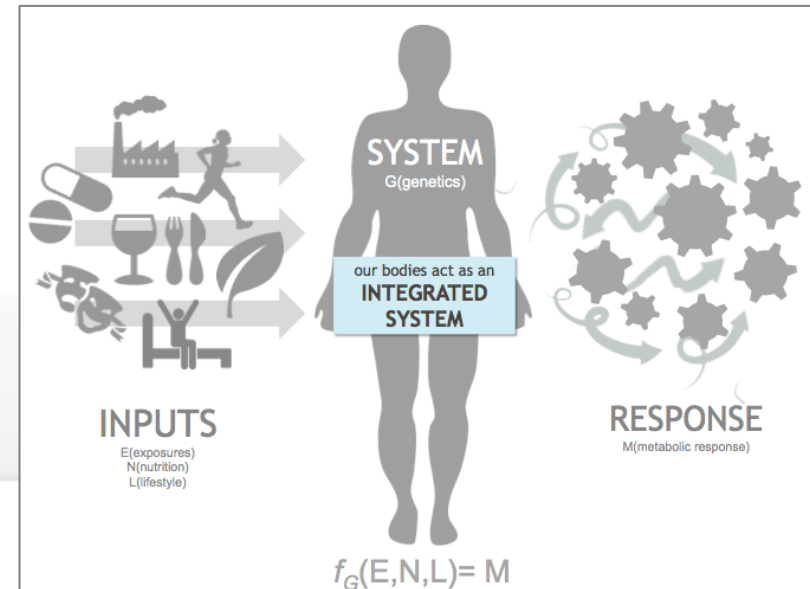
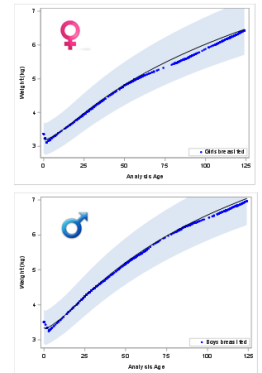
- Feed R&D

Substantiate results, new hypotheses



Learn more about our data

- Power to detect subtle differences
- Broader picture, beyond single study
- New dimensions, e.g ingredients



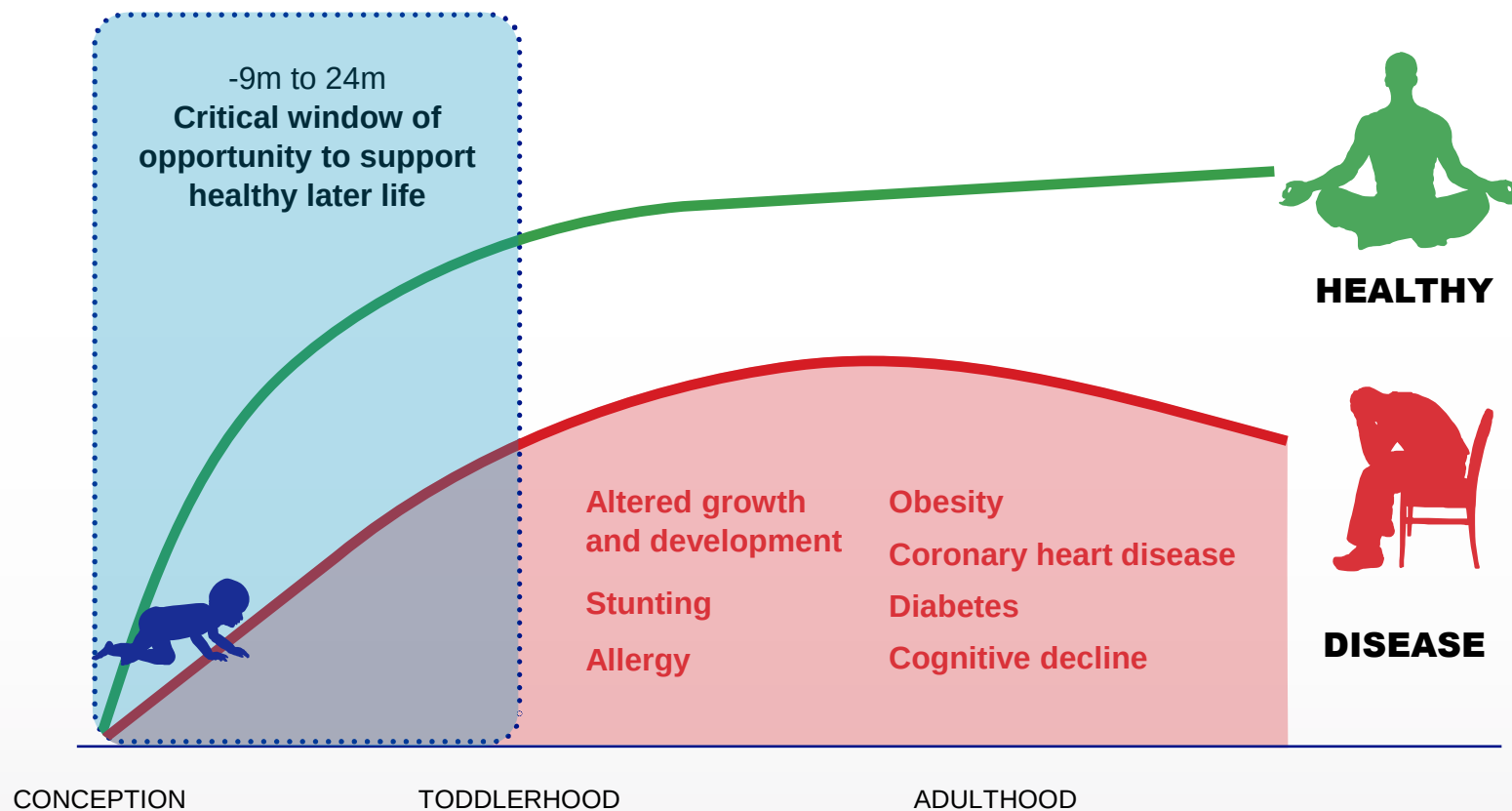
PART OF DANONE – FOUR DIVISIONS



“Bring health through food
to as many people as possible”

THE FIRST 1000 DAYS

AN IMPACT ON LATER LIFE HEALTH



A BABY IS NOT A SMALL ADULT

THE ORGANS OF A NEWBORN BABY STILL NEED TO DEVELOP

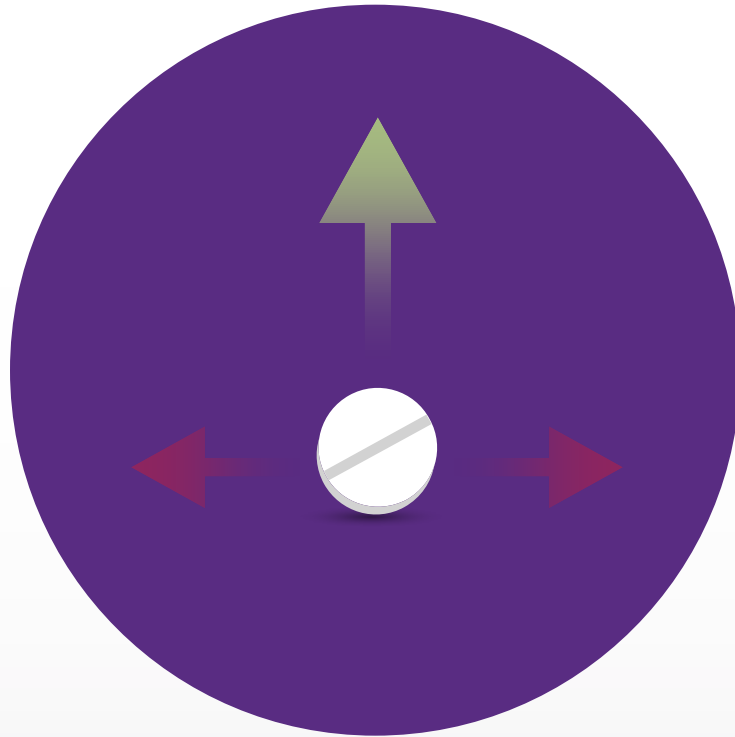
Immaturity of:



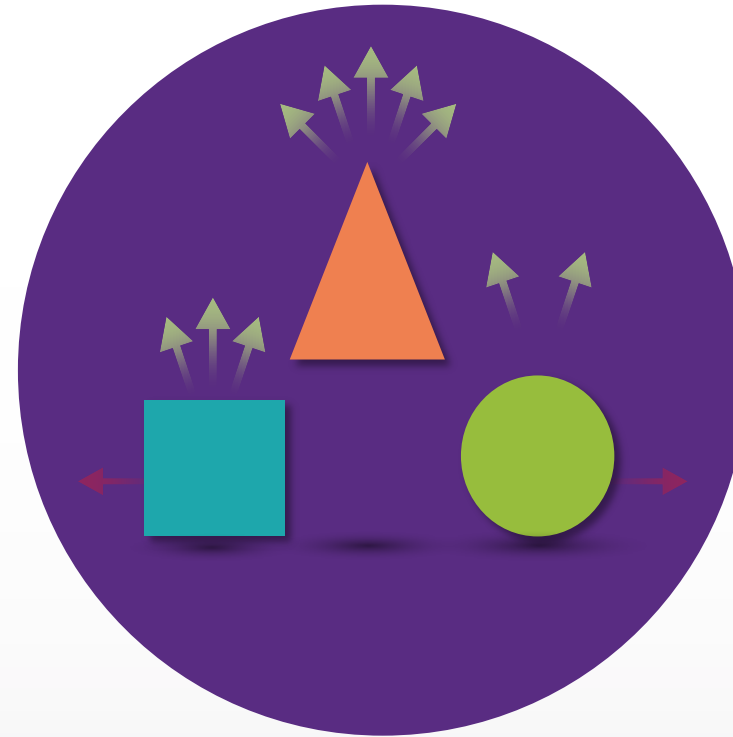
THE SCIENCE OF NUTRITION

DIFFERENCES IN ACTION REQUIRE SPECIFIC RESEARCH APPROACH

**FOCUSED LARGE EFFECT ON
SINGLE TARGET**



**CUMULATIVE EFFECTS ON
MULTIPLE TARGETS**



QUOTES FROM PROGRAMMERS AND STATISTICIANS

Analyses of secondary parameters leads to surprising and very interesting patterns!

Many exploratory endpoints and analyses

No NDAs, thus no external incentive to standardize clinical trial data across studies

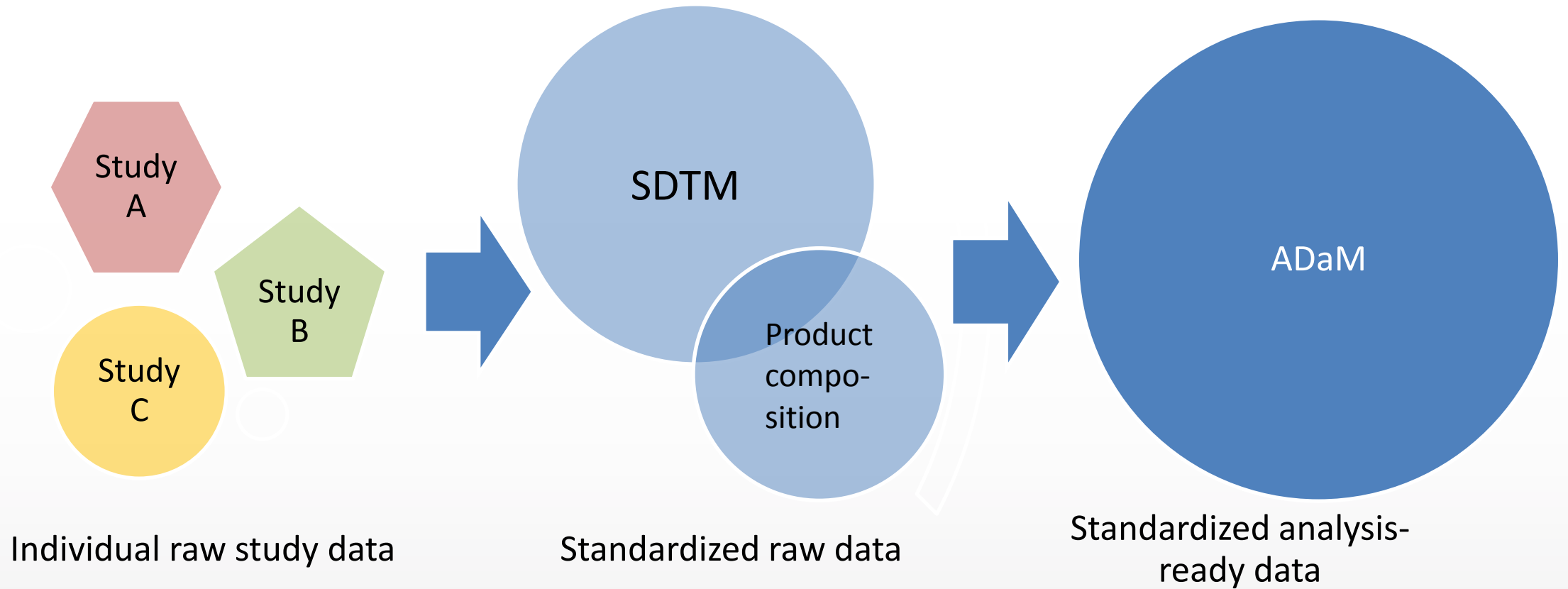
Long-term follow-up in quite a number of studies

Open-minded culture, with standards and processes still developing and maturing

Not the standard Phase I-IV trials, and fewer trials for one product than in Pharma, but each trial investigates many outcomes



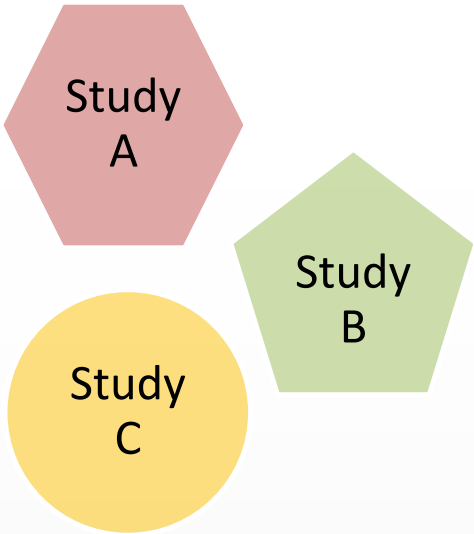
DATA STANDARDIZATION AND INTEGRATION



GROWTH DATA – VITAL SIGNS (VS)

randomno weightinf_W4 lengthinf_W4 headcirc_W4 weighinf_W8 ...

	randomno	weightinf_W4	lengthinf_W4	headcirc_W4	weightinf_W8	lengthinf_W8	headcirc_W8	weightinf_W13	lengthinf_W13	headcirc_W13
▶ 1	1	5120	56	38,3	5800	58,7	39,2	6680	62,2	40,6



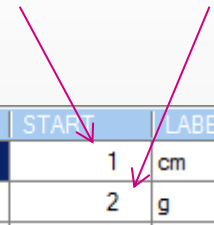
VISIT USUBJID VSTEST VSORRES VSORRESU

	VISIT	USUBJID	VSTEST	VSORRES	VSORRESU
▶ 1	Visit 1	Dig.1.C/A-056-41-2201	Weight	3540	grams
2	Visit 1	Dig.1.C/A-056-41-2201	Height	50.0	cm
3	Visit 1	Dig.1.C/A-056-41-2201	Head circumference	33.0	cm
4	Visit 2	Dig.1.C/A-056-41-2201	Weight	4100	grams
5	Visit 2	Dig.1.C/A-056-41-2201	Height	54.5	cm
6	Visit 2	Dig.1.C/A-056-41-2201	Head circumference	36.0	cm
7	Visit 3	Dig.1.C/A-056-41-2201	Weight	4780	grams
8	Visit 3	Dig.1.C/A-056-41-2201	Height	56.0	cm
9	Visit 3	Dig.1.C/A-056-41-2201	Head circumference	37.5	cm

PID VISIT HEIG HEIGU WEIG WEIGU

	PID	VISIT	HEIG	HEIGU	WEIG	WEIGU
▶ 1	2001	2	50	1	3240	2
2	2001	8	55.5	1	5100	2
3	2001	16	65	1	6400	2

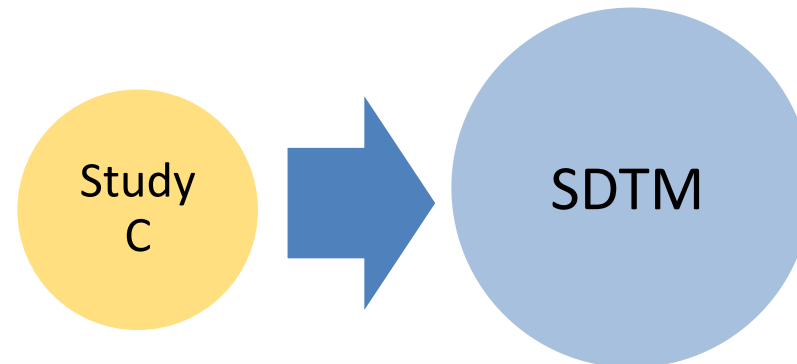
FMTNAME	START	LABEL
UNITSF	1	cm
UNITSF	2	g



GROWTH DATA – VITAL SIGNS (VS)

Codelist Name	CDISC Submission Value	CDISC Synonym(s)	CDISC Definition
Vital Signs Test Code	VSTESTCD	Vital Signs Test Code	The name given to the test that analyzes a particular set of vital signs including temperature, respiratory rate, heart beat (pulse), and blood pressure. (NCI)
Vital Signs Test Code	WEIGHT	Weight	The vertical force exerted by a mass as a result of gravity. (NCI)
Vital Signs Test Code	HDCIRC	Head Circumference	A circumferential measurement of the head at the widest point.
Vital Signs Test Code	BODLNTH	Body Length	The linear extent in space from one end of the body to the other end, or the extent of the body from beginning to end.

Codelist Name	CDISC Submission Value	CDISC Synonym(s)	CDISC Definition
Units for Vital Signs Results	VSRESU	Units for Vital Signs Results	The unit used to record and describe the result of a test investigating a vital sign. (NCI)
Units for Vital Signs Results	cm	Centimeter	A basic unit of length equal to one hundredth of a meter or approximately 0.393700787 inch.
Units for Vital Signs Results	g	Gram	A metric unit of mass equal to one one thousandth of a kilogram. (NCI)



VISIT USUBJID VSTEST VSORRES VSORRESU

	VISIT	USUBJID	VSTEST	VSORRES	VSORRESU
1	Visit 1	DIG.1.C/A-056-41-2201	Weight	3540	grams
2	Visit 1	DIG.1.C/A-056-41-2201	Height	50.0	cm
3	Visit 1	DIG.1.C/A-056-41-2201	Head circumference	33.0	cm
4	Visit 2	DIG.1.C/A-056-41-2201	Weight	4100	grams
5	Visit 2	DIG.1.C/A-056-41-2201	Height	54.5	cm
6	Visit 2	DIG.1.C/A-056-41-2201	Head circumference	36.0	cm
7	Visit 3	DIG.1.C/A-056-41-2201	Weight	4780	grams
8	Visit 3	DIG.1.C/A-056-41-2201	Height	56.0	cm
9	Visit 3	DIG.1.C/A-056-41-2201	Head circumference	37.5	cm



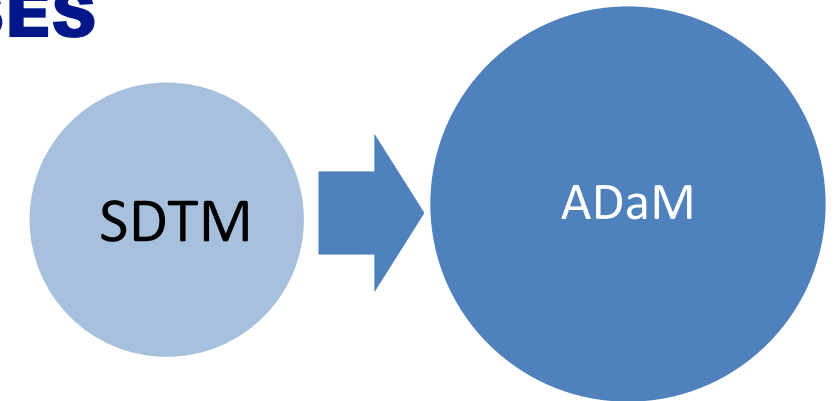
DOMAIN USUBJID VSTESTCD VSTEST VSORRES VSORRESU VISIT

	DOMAIN	USUBJID	VSTESTCD	VSTEST	VSORRES	VSORRESU	VISIT
1	VS	DIG.1.C/A-056-041-2201	BODLNTH	Body Length	50.0	cm	Visit 1
2	VS	DIG.1.C/A-056-041-2201	HDCIRC	Head Circumference	33.0	cm	Visit 1
3	VS	DIG.1.C/A-056-041-2201	WEIGHT	Weight	3540	g	Visit 1
4	VS	DIG.1.C/A-056-041-2201	BODLNTH	Body Length	54.5	cm	Visit 2
5	VS	DIG.1.C/A-056-041-2201	HDCIRC	Head Circumference	36.0	cm	Visit 2
6	VS	DIG.1.C/A-056-041-2201	WEIGHT	Weight	4100	g	Visit 2
7	VS	DIG.1.C/A-056-041-2201	BODLNTH	Body Length	56.0	cm	Visit 3
8	VS	DIG.1.C/A-056-041-2201	HDCIRC	Head Circumference	37.5	cm	Visit 3
9	VS	DIG.1.C/A-056-041-2201	WEIGHT	Weight	4780	g	Visit 3

ONE STEP AWAY FROM GROWTH ANALYSES

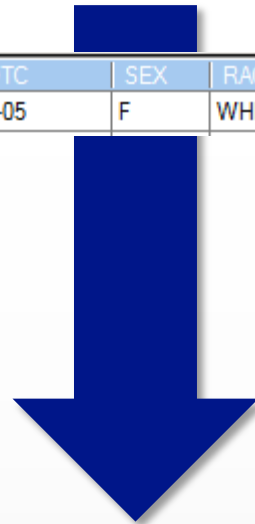
SDTM.VS

	DOMAIN	USUBJID	VSTESTCD	VSTEST	VSORRES	VSORRESU	VISIT	VSDTC
▶ 1	VS	DIG.1.C/A-056-041-2201	WEIGHT	Weight	3540	g	Visit 1	2011-12-06
2	VS	DIG.1.C/A-056-041-2201	WEIGHT	Weight	4100	g	Visit 2	2012-01-04
3	VS	DIG.1.C/A-056-041-2201	WEIGHT	Weight	4780	g	Visit 3	2012-02-01



SDTM.DM

	DOMAIN	USUBJID	BRTHDTC	SEX	RACE
▶ 1	DM	DIG.1.C/A-056-041-2201	2011-12-05	F	WHITE

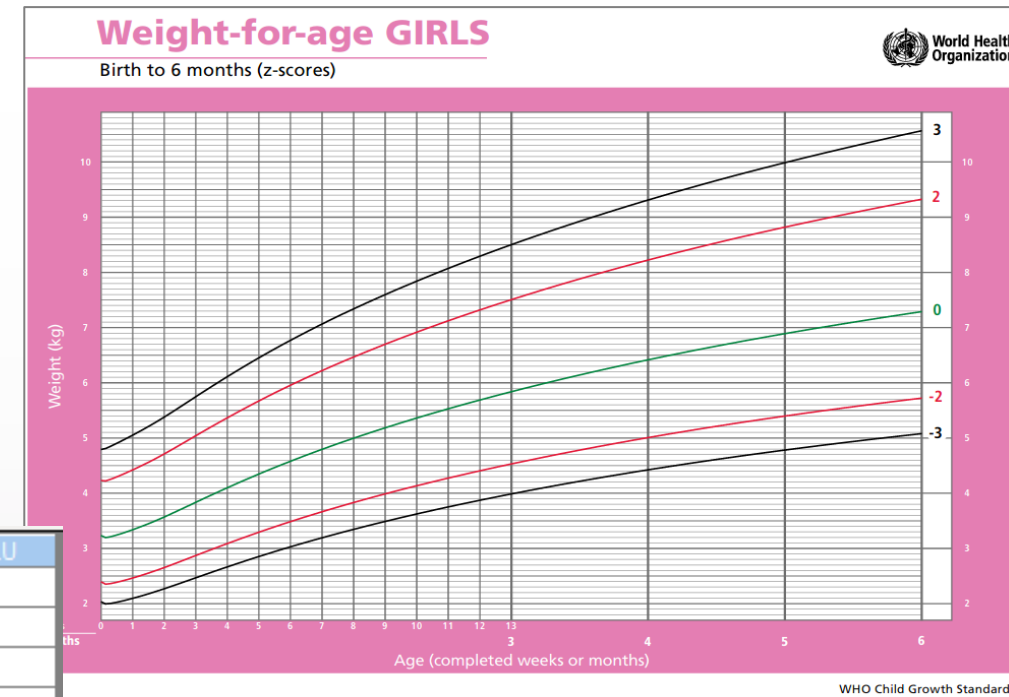


WHO reference data

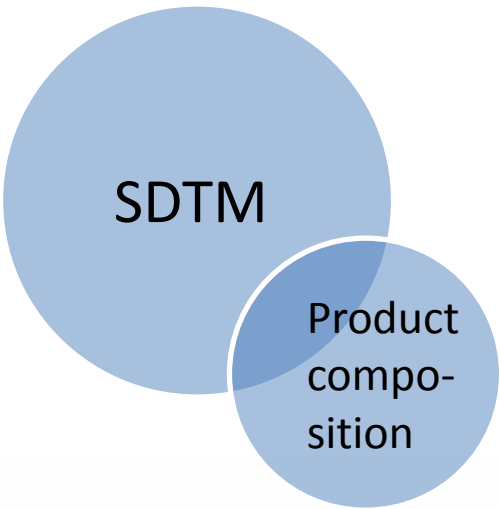
Age	-1SD	Median	+1SD
0	2.794	3.232	3.711
1	2.752	3.196	3.685
30	3.630	4.172	4.778
58	4.429	5.051	5.750

ADaM.ADVS

	USUBJID	AAGE	AAGEU	VISIT	AVISIT	PARAM	PARAMCD	AVAL	AVALU
60003	DIG.1.C/A-056-041-2201	1	DAYS	Visit 1	Baseline	Weight	WEIGHT	3,54	kg
60004	DIG.1.C/A-056-041-2201	1	DAYS		Baseline	Weight-for-age z-score	ZWEI	0,71	
60005	DIG.1.C/A-056-041-2201	30	DAYS	Visit 2	Month 1	Weight	WEIGHT	4,1	kg
60006	DIG.1.C/A-056-041-2201	30	DAYS		Month 1	Weight-for-age z-score	ZWEI	-0,13	
60007	DIG.1.C/A-056-041-2201	58	DAYS	Visit 3	Month 2	Weight	WEIGHT	4,78	kg
60008	DIG.1.C/A-056-041-2201	58	DAYS		Month 2	Weight-for-age z-score	ZWEI	-0,42	



PRODUCT COMPOSITION AS DATA



Product Characteristic	Product Characteristic Category	Product Characteristic Description
PRDCHAR	CHARCAT	CHARDESC
PREB10	Prebiotics	Prebiotics lcFOS
PREB20	Prebiotics	Prebiotics scGOS
PROT1.8	Protein	Protein at 1.8g/100kcal

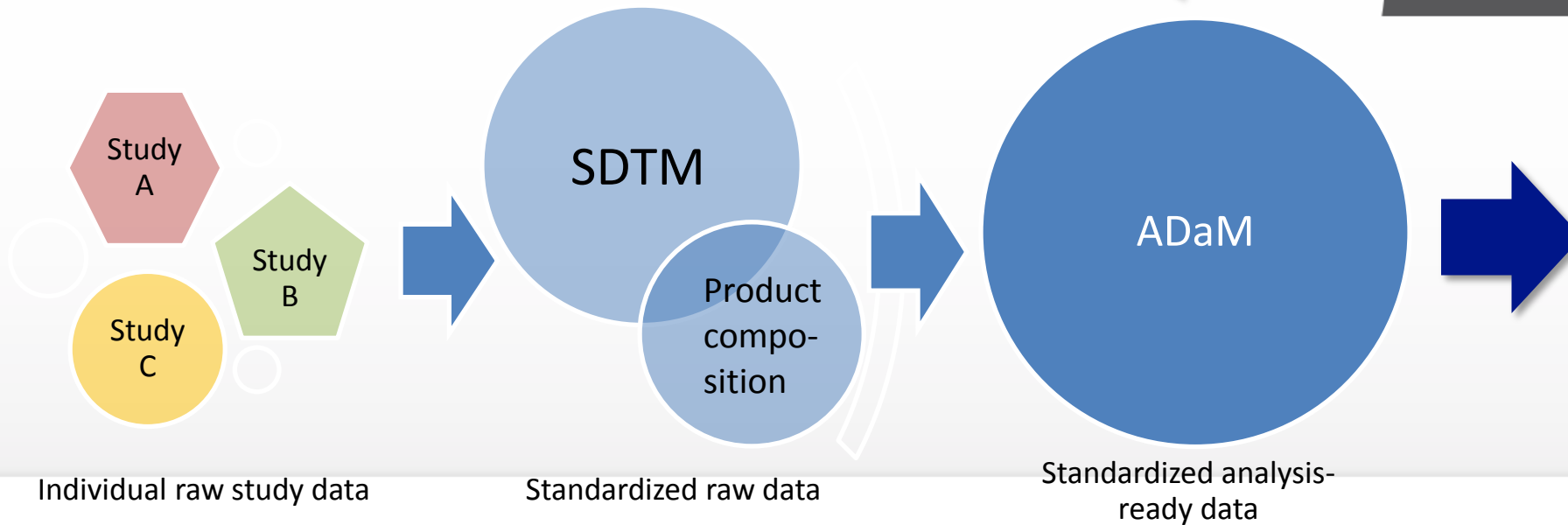
Unique Product Identifier	Unique Product Full Description	Product Characteristics Space Delimited List From CT
UPRODID	UPROD	PRDCHAR
CFLP00	Low Protein Cow's Milk Infant Formula with a protein content of 1.8 g/100kcal	PROT1.8
CFLP10	Modified Low Protein Cow's Milk Infant Formula with a protein content of 1.8 g/100kcal with ferments and scGOS/lcFOS	PROT1.8 PREB10 PREB20 FERM

Planned treatment (ARM/TRT01P)

Analysis Record Flag (ANLzzFL)

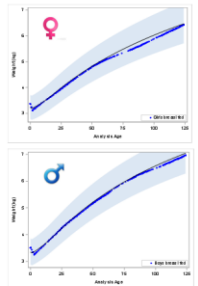
'ANALYSIS ENGINE'

What is the effect of ingredient xx on growth across all studies? (and I want a p-value)



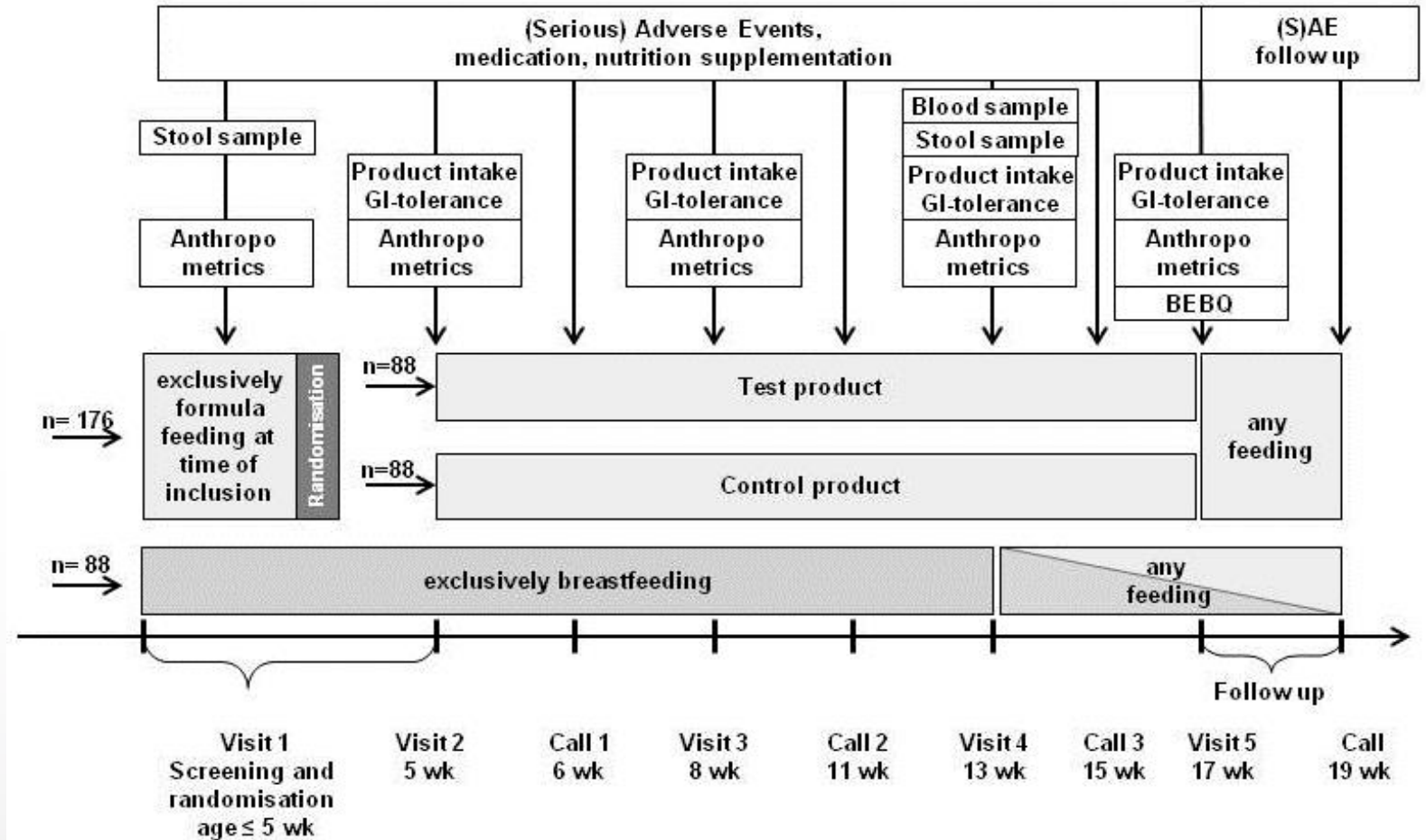
Learn more about our data

- Power to detect subtle differences
- Broader picture, beyond single study
- New dimensions, e.g ingredients



SPECIFIC DOMAINS IN NUTRITION CLINICAL STUDIES

- ❑ **Many common domains:**
DM, AE, CM, MH, LB
- ❑ **Typical:**
 - GI-tolerance
 - Microbiota data
 - Food intake
 - Parent data



GI ?

GLYCEMIC INDEX CHART									
Low Glycemic (55 or Below)					High Glycemic (70 or Higher)				
									
SNACKS	G.I.	STARCH	G.I.	VEGETABLES	G.I.	FRUITS	G.I.	DAIRY	G.I.
Pizza	33	Bagel, Plain	33	Broccoli	10	Cherries	22	Yogurt, Plain	14
Chocolate Bar	49	White Rice	38	Pepper	10	Apple	38	Yogurt, Low Fat	14
Pound Cake	54	White Spaghetti	38	Lettuce	10	Orange	43	Whole Milk	30
Popcorn	55	Sweet Potato	44	Mushrooms	10	Grapes	46	Soy Milk	31
Energy Bar	58	White Bread	49	Onions	10	Kiwi	52	Skim Milk	32
Soda	72	Brown Rice	55	Green Peas	48	Banana	56	Chocolate Milk	35
Doughnut	76	Pancakes	67	Carrots	49	Pineapple	66	Yogurt, Fruit	36
Jelly Beans	80	Wheat Bread	80	Beets	64	Watermelon	72	Custard	43
Pretzels	83	Baked Potato	85	Onions	75	Dates	103	Ice Cream	60

Glycemic Index values obtained from www.lowglycemicdiet.com, www.nutritiondata.com and www.diabetesnet.com



GI



GI TOLERANCE

CRF example

Visit 1 – Baseline

Gastrointestinal tolerance

Does the child / the subject suffer from:

Nausea	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Vomiting	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Burping	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Flatulence	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Diarrhoea	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Constipation	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Colic (cramps)	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Regurgitation	☐ Absent	☐ Mild	☐ Moderate	☐ Severe

Average number of stools per day:

What is the average consistency of the stools?

- ☐ Watery
- ☐ Soft, puddinglike
- ☐ Soft, formed
- ☐ Dry, formed
- ☐ Dry, hard pellets

GI TOLERANCE

CRF example

Visit 1 – Baseline

Gastrointestinal tolerance

Does the child / the subject suffer from:

Nausea	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Vomiting	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Burping	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Flatulence	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Diarrhoea	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Constipation	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Colic (cramps)	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Regurgitation	☐ Absent	☐ Mild	☐ Moderate	☐ Severe

Average number of stools per day:

What is the average consistency of the stools?

- ☐ Watery
- ☐ Soft, puddinglike
- ☐ Soft, formed
- ☐ Dry, formed
- ☐ Dry, hard pellets

GENERAL OBSERVATION CLASSES (SDTM-IG 2.1)

- The **Interventions** class captures investigational, therapeutic and other treatments that are administered to the subject (with some actual or expected physiological effect) either as specified by the study protocol (e.g., exposure to study drug), coincident with the study assessment period (e.g., concomitant medications), or self-administered by the subject (such as use of alcohol, tobacco, or caffeine).
- The **Events** class captures planned protocol milestones such as randomization and study completion, and occurrences, conditions, or incidents independent of planned study evaluations occurring during the trial (e.g., adverse events) or prior to the trial (e.g., medical history).
- The **Findings** class captures the observations resulting from planned evaluations to address specific tests or questions such as laboratory tests, ECG testing, and questions listed on questionnaires.

GI TOLERANCE

CRF example

Visit 1 – Baseline

Gastrointestinal tolerance

Does the child / the subject suffer from:

CETERM	CEPRESP='Y'	CEOCCUR='Y'			
		CESEV			
Nausea	☒ Absent	☐ Mild	☐ Moderate	☐ Severe	
Vomiting	☐ Absent	☒ Mild	☐ Moderate	☐ Severe	
Burping	☐ Absent	☐ Mild	☒ Moderate	☐ Severe	
Flatulence	☐ Absent	☒ Mild	☐ Moderate	☐ Severe	
Diarrhoea	☒ Absent	☐ Mild	☐ Moderate	☐ Severe	
Constipation	☒ Absent	☐ Mild	☐ Moderate	☐ Severe	
Colic (cramps)	☒ Absent	☐ Mild	☐ Moderate	☐ Severe	
Regurgitation	☒ Absent	☐ Mild	☐ Moderate	☐ Severe	

CEOCCUR='N'

DOMAIN	USUBJID	CETERM	CECAT	CEPRESP	CEOCCUR	CESEV
CE	XXX-123-001	NAUSEA	GASTROINTESTINAL TOLERANCE	Y	N	
CE	XXX-123-001	VOMITING	GASTROINTESTINAL TOLERANCE	Y	Y	MILD
CE	XXX-123-001	BURPING	GASTROINTESTINAL TOLERANCE	Y	Y	MODERATE
CE	XXX-123-001	FLATULENCE	GASTROINTESTINAL TOLERANCE	Y	Y	MILD

GI TOLERANCE

CRF example

Visit 1 – Baseline

Gastrointestinal tolerance

Does the child / the subject suffer from:

Nausea	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Vomiting	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Burping	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Flatulence	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Diarrhoea	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Constipation	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Colic (cramps)	☐ Absent	☐ Mild	☐ Moderate	☐ Severe
Regurgitation	☐ Absent	☐ Mild	☐ Moderate	☐ Severe

Average number of stools per day:

What is the average consistency of the stools?

- ☐ Watery
- ☐ Soft, puddinglike
- ☐ Soft, formed
- ☐ Dry, formed
- ☐ Dry, hard pellets

FINDINGS ABOUT (FA)

- A dataset used to capture the findings about an event or intervention that cannot be represented within an event or intervention record or as a supplemental qualifier.
- Criteria 1) data that do not describe an Event or Intervention as a whole, 2) data (“about” an Event or Intervention) having Qualifiers that can be represented in Findings variables (e.g., units, method), and 3) data (“about” an Event or Intervention) that indicate the occurrence of related symptoms or therapies.
- FAOBJ: Used to describe the object or focal point of the findings observation that is represented by FATEST. Examples: an event such as VOMIT where the volume of Vomit is being measured by a VOLUME test.

6 Domain Models Based on the General Observation Classes

6.4 Findings About Events or Interventions

Findings About (FA)

FA – Description/Overview for Findings About Domain Model

A dataset used to capture the findings about an event or intervention that cannot be represented within an event or intervention record or as a supplemental qualifier.

FA – Specification for Findings About Domain Model

fa.xpt, Findings About Events or Interventions — Findings Sub-Class, Version 3.1.2. One record per finding, per object, per time point, per visit per subject Tabulation

Variable Name	Variable Label	Type	Controlled Terms, Code list or Format	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
DOMAIN	Domain Abbreviation	Char	FA	Identifier	Two-character abbreviation for the domain.	Req
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	Req
FASEQ	Sequence Number	Num		Identifier	Sequence Number given to ensure uniqueness of subject records within a domain. May be any valid number.	Req
FAGRPID	Group ID	Char		Identifier	Used to tie together a block of related records in a single domain for a subject.	Perm
FASPID	Sponsor-Defined Identifier	Char		Identifier	Sponsor-defined reference number. Perhaps pre-printed on the CRF as an explicit line identifier or defined in the sponsor's operational database. Example: Line number on a CRF.	Perm
FATESTCD	Findings About Test Short Name	Char	(FATESTCD)	Topic	Short name of the measurement, test, or examination described in FATEST. It can be used as a column name when converting a dataset from a vertical to a horizontal format. The value in FATESTCD cannot be longer than 8 characters, nor can it start with a number (e.g. "1 TEST"). FATESTCD cannot contain characters other than letters, numbers, or underscores. Example: SEV, OCCUR.	Req

© 2013 Clinical Data Interchange Standards Consortium, Inc. All rights reserved.
Final

FA - Page 1
November 26, 2013

GI TOLERANCE

Visit 1 – Baseline

Gastrointestinal tolerance

DOMAIN	USUBJID	FATESTCD	FATEST	FAOBJ	FAORRES	FAORRESU
FA	YYY-321-001	A_COUNT	Average Count	STOOL	3	1/DAY
FA	YYY-321-001	A_CONSIS	Average Consistency	STOOL	WATERY	

FATEST='Average Count'

FAOBJ='STOOL'

Average number of stools per day:

3

FATEST='Average consistency'

FAORRES

What is the average consistency of the stools?

☒ Watery

☐ Soft, puddinglike

☐ Soft, formed

☐ Dry, formed

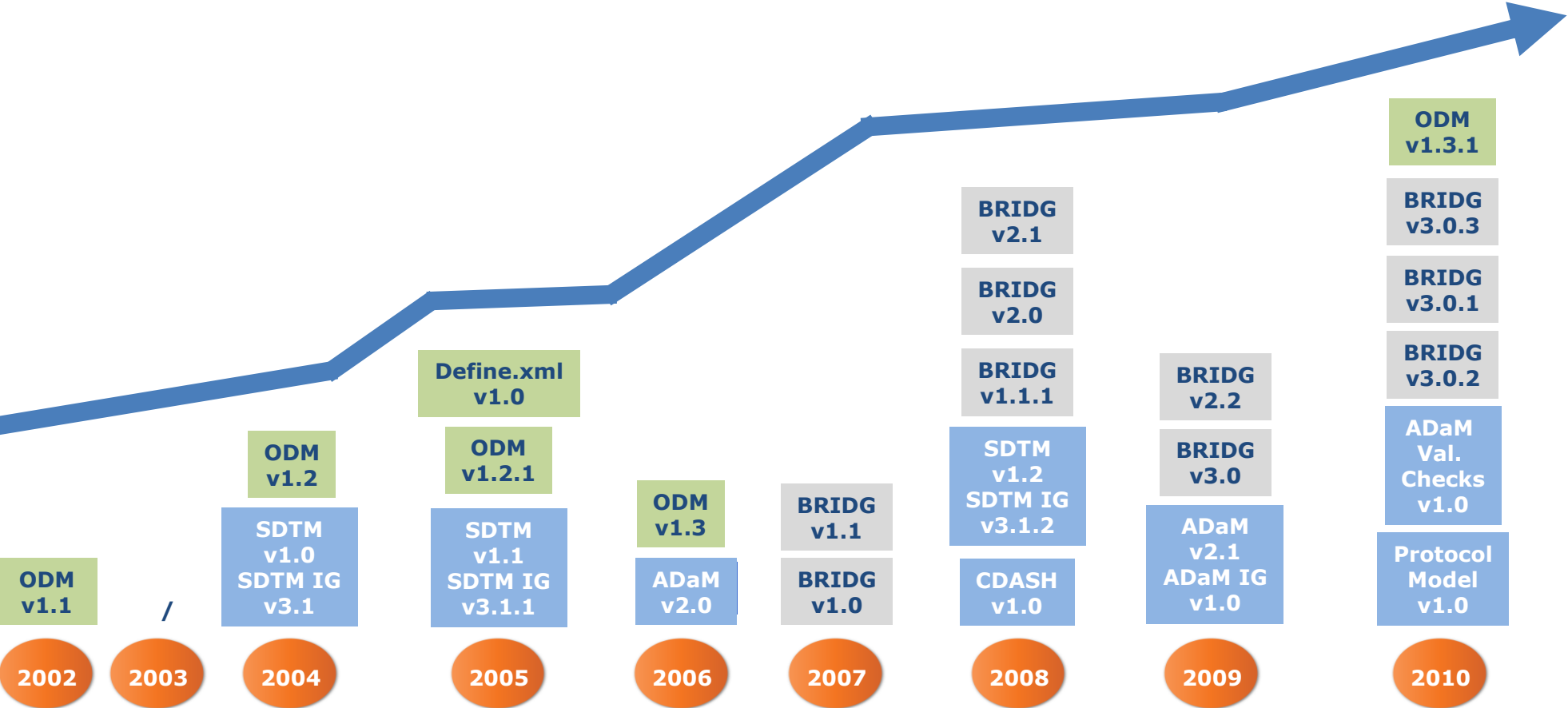
☐ Dry, hard pellets

Updates over time

(MAY 2016)

Kindly provided by:
Peter van Reusel
Business & Decision Life Sciences

- Content Standards
- Technical Standards
- Semantics
- Therapeutic Areas

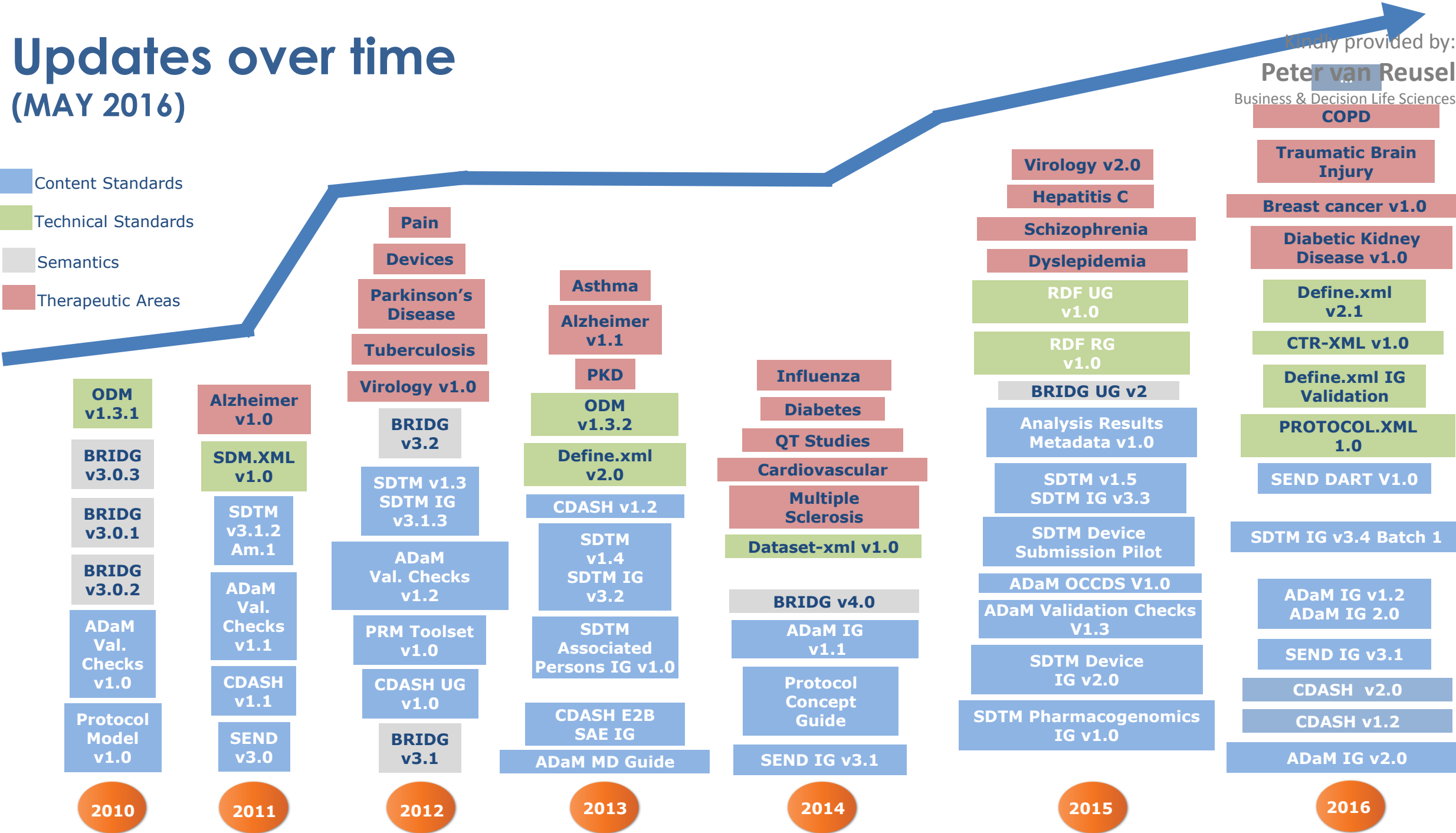


Updates over time

(MAY 2016)

Kindly provided by:
Peter van Reusel
Business & Decision Life Sciences

- Content Standards
- Technical Standards
- Semantics
- Therapeutic Areas



WHAT DOES THE FUTURE LOOK LIKE?

Leverage value and increase capabilities

- Build on cross-functional collaborations and R&D data strategy
- Streamline with upstream standardization (CDASH/SDTM)
- Enable controlled data sharing and usage for Nutricia Research R&D community
- Incorporate & explore new sources of data

There's more to come...





THANK YOU

**Sophie Swinkels
Marieke Abrahamse
Lieke Egbers
Daniela Cianci
John Roth
Remco van der Meer
Michael Auld
Lieke Gijsbers
Lennert van der Zee
Jules van der Zalm**

**NUTRICIA
RESEARCH**

