

Data Collected: What data has been collected when and on whom

Mary Pettinger, WHI CCC
Fred Hutch, Seattle WA
September 9, 2016



Objectives

- Familiarize investigators with the data collected in all components and phases of WHI (1993-2016)
- Provide information to enable users to get started on their own to propose a paper, ancillary study and/or conduct analyses

Plan to Cover:

What

- Questionnaires
- Clinical exams
- Outcomes
- Biospecimens

When

- Screening
- WHI (thru 2005)
- Extension Studies
 - I (2005-2010)
 - II (2010-2015)
 - III (2015-present)

Whom

- All
- Subsamples
- Subsets

Motivation for Data Collection: CT

- Designated primary and secondary clinical outcomes in each CT
- Interest in intermediate outcomes (i.e. lipid levels in HT)
- Trial monitoring, safety, adherence
- Potential effect modifiers

Motivation for Data Collection: OS

- to examine the associations of known or putative risk factors (and protective factors) to disease
- to find new risk factors and/or biomarkers using the stored biologic samples and data as a resource
- to examine the association of change in individual characteristics on disease and mortality

Abbreviations:

- CT = Clinical Trial
- OS = Observational Study
- HT/H = Hormone Trial
(includes both E-alone and E+P)
- DM/D = Dietary Modification Trial
- CaD = Calcium and Vitamin D Trial
- BMD/BD = Bone Density sites
- MRC = Medical Record Cohort
- SRC = Self Report Cohort
- % = Subset or subsample



What Data and When

General Categories

- Demographics
- Medical and physical measurements
- Medical and family History
- Diet
- Psychosocial and personal habits
- Medication and supplement use
- Outcomes
- Specimen results

Form #	Form Name	Screening CT and OS				CT																	OS												
		SV 0	SV 1	SV 2	SV 3	4-6 wk	6 m	1 Yr	4 wk	6 m	2 Yr	6 m	3 Yr	6 m	4 Yr	6 m	5 Yr	6 m	6 Yr	6 m	7 Yr	6 M	8 Yr	6 m	9 Yr	Close Out	An-nual	3 Yr	6 Yr	9 Yr					
90	Functional Status				%HD			%					%					%					%												
92	Pap			H								H						H					H												
100	Blood Collection		X					X					%					%					%									X			
101	Urine Collection		BD					BD				BD											BD							BD		BD			
143	OS Follow-up (Year 3)																														X				
144	OS Follow-up (Year 4) ¹																																		
145	OS Follow-up (Year 5) ¹																																		
146	OS Follow-up (Year 6)																																X		
147	OS Follow-up (Year 7) ¹																																		
148	OS Follow-up (Year 8) ¹																																		
149	Supplement to OS Follow-up ¹																																		

Key: X = All Participants
 D = DM
 H = HRT
 C = CaD
 O = OS
 % = Percentage (subsample) of participants
 BD = Bone Density sites

Screening

Form		Screening CT and OS			
#	Name	SV 0	SV 1	SV 2	SV 3
2	Eligibility Screen	X			
4	HRT Washout		H		
20	Personal Information		X	X	X
30	Medical History		X		
34	Personal Habits				X

OS Follow-up – WHI (-2005)

Form		OS			
#	Name	Annual	3 Yr	6 Yr	9 Yr
33	Medical History Update	X	X	X	X
38	Daily Life		X		
44	Current Medications		X		
45	Current Supplements		X		
60	Food Questionnaire (FFQ)		X		
80	Physical Measurements		X		

OS Follow-up – data items

- Annual: smoking, hormone use
- Repeated: alcohol (yrs 1, 3, 6), physical activity
- One-time only: dental health (yr 5), sun exposure & artificial sweetener use (yr 4), insecticide exposure (yr 1)

Extension Study I (2005-2010)

Form #	Form Name	1 Yr	2 Yr	3 Yr	4 Yr	5 Yr
-	Re-consent and personal information update	X				
33	Medical History Update	X	X	X	X	X
85	Mammogram	H	H			
150	Hormone Use Update	H	H	H	H	H
151	QOL, Activities of Daily Living (ADL)	X	X	X	X	X
153	Medication and supplement inventory				X	
-	24 Hour Dietary Recall	%D		%D		

Key

X = All Participants

D = DM

H = HT

% = Percentage (subsample) of participants

Extension Study II (2010-2015)

Form #	Form Name	1 Yr	2 Yr	3 Yr	4 Yr	5 Yr
-	Re-consent and personal information update	X				
33	Medical History Update	X	X	X	X	X
151	QOL, Activities of Daily Living (ADL)	X		X	X	X
155	Lifestyle Qx (includes QOL, ADL)		X			
153	Medication and supplement inventory			MRC		
156	Supplemental Qx				X	
157	Supplemental Qx 2014-2015					X

Key

X = All Participants

MRC = Medical Record Cohort

Extension Study III (2015-)

Form #	Form Name	1 Yr	2 Yr	3 Yr	4 Yr	5 Yr
33	Medical History Update	X	X	X	X	X
151	QOL, Activities of Daily Living (ADL)	X	X	X	X	X
521	Physical Activity Questionnaire	X				
158	Supplemental Qx 2016-2017		X			

Key

X = All Participants

Extension Study Data Items

Shift of focus to aging related questions:

- ADL, physical functioning
- Vaccines for shingles, pneumonia, flu
- Living assistance
- Social support
- Health care and advanced planning

Key Forms

- Baseline:

Form 2 & 30 - medical history

Form 20 – SES

Form 2 & 41 – race and ethnicity

Form 34 – personal habits

Form 80 – physical characteristics (height, weight, BMI, blood pressure)

Form 37 – QOL, psychosocial

Form 43 – Hormone use history

Form 44 & 45 – medications & supplements

Key Forms, continued

- Follow-up:

Form 35 – personal habits update

Form 38 – psychosocial update

Form 80 – physical characteristics (height, weight, BMI, blood pressure)

- Extension follow-up:

Form 151 – QOL, ADL

Form Level Datasets

- Include responses to the individual questions
- Plus, computed variables
 - Ex: Form 34 asks about smoking history – calculated ‘pack years’
 - Form 37 has ~20 additional behavioral/psychosocial constructed variables

Data Collected on Which Participants

- All, ex: Personal habits at screening, medical history during follow-up
- Subset, ex: clinical exams (HT only), bone density (3 centers), cognitive function (HT, age > 65)
- Subsample => implies a sampling scheme, defined at the beginning of WHI
ex: FFQ after year1, blood analytes.

WHI Subsample Overview

Subsample	Design	Component	%	Years	Detail	Required operations
Blood	Cohort	HT	8.6%	0,1,3,6,9	Stratified by CC, hyst (HT), race; oversampling of minorities w/ goal of equity across clinics.	Blood collection at yrs 0, 1, 3,6, 9 (Form 100); blood analyzed prospectively at yrs 0, 1, 3, 6, 9
		DM (CT)	4.3% (~6%)			
FFQ	One-time		30%	2	Stratified by CC	Designed so that each woman is scheduled to complete an FFQ every 3 years after year 2
	Rotating, cross-sectional	DM	33%	Annually after yr 2		
4DFR	Cohort	DM	4.6%	0, 1	Overlap with blood subsample in DM	4DFRs collected, documented and analyzed at yrs 0, 1 on subsample
24HR Recall	Repeated, cross-sectional	DM	1%	Annually	Selected annually	24HR administered centrally
	Cohort	DM	4.6%	3, 6, 9	Same subsample as 4DFR	Note: 24HR replaced the 4DFR after year 1
Endometrial Aspiration	Cohort	HT w/ uterus	5%	3, 6, 9	Stratified by CC	Endometrial aspiration required at years 3, 6, 9 (Form 82)
Cognitive Function	Cohort	HT > 65 yrs old	25%	0,1,3,6,9	Stratified by CC	Form 39 at years 1,3,6,9
Functional Status	Cohort	HT > 65 yrs old DM > 65 yrs old	25% 25%	0,1,3,6,9	Stratified by CC, selected independently of Cognitive Function sample	Form 90 at years 1,3,6,9
Daily Life	Cohort	HT	8.6%	3,6,9	Same as blood subsample	Form 38 at years 3,6,9
		DM	4.3%			
OS-MPS	One-time	OS	1%	0	Stratified by CC, age, race/ethnicity, subject to consent	Repeat selected baseline measures at 3 months, including blood sample

Core Blood Subsample (W1)

- Important: blood was collected on all CT & OS participants at screening, all CT at year 1 but only a subsample at years 3, 6 and 9; and all OS at year 3
- 20 Analytes* measured on only a subsample, oversampled for minorities
 - ~1% of OS samples at screening
 - ~8.6% of HT, and ~4.3% of DM samples at screening, and years 1, 3, and 6

*lipids, insulin & glucose, micronutrients and clotting factors

Race/Ethnicity of Core Blood Subsample

CT Subsample(n=4,427)

49% White

26% Black

13% Hispanic

3% American Indian

8% Asian/PI

2% Unknown

Entire CT (n=68,132)

81% White

10% Black

4% Hispanic

0.4% American Indian

2% Asian/PI

1% Unknown

Adjust for sampling in analyses

- Direct adjustment when reporting Mean \pm SD, or frequencies
- Inclusion of race/ethnicity covariate in modeling
- Or a type of weighting for sampling distribution

Additional Subsets – Post WHI

Starting with Extension II (2010-2015), participants split into:

- 1) Medical Records Cohort (MRC) – All HT, Black and Hispanic participants
CVD, cancer and death outcomes adjudicated
- 2) Self-Report Cohort (SRC) – non MRC
Cancer outcomes adjudicated; otherwise self-report only

More Subsets

- LLS: Long Life Study
 - OPACH: Objective Physical Activity and Cardiovascular Health
- LILAC: Life and Longevity After Cancer
- New Randomized Trials
 - WHISH: Women's Health Initiative Strong and Healthy
 - COSMOS: COcoa Supplement and Multivitamin Outcomes Study

Biospecimen Collection

- Baseline—All participants
 - Serum
 - Plasma (citrate & EDTA)
 - Buffy coat
 - RBC
 - Urine (BMD subsample—3 sites only)
- Follow-up: Blood
 - CT: 100% year 1, 6% years 3,6,9 (no DNA)
 - OS: 100% year 3
- Follow-up: Urine
 - BMD subsample: 100% years 1,3,6,9 (no blood)
- LLS (Extension Study II) – serum, plasma, RBC, DNA, RNA
- Details:
<https://www.whi.org/about/SitePages/Biospecimen%20Collection.aspx>

Also collected at time of specimen draw

- Status of Participant
 - Hours since last ate
 - Vigorous activity in previous 8 hours
 - Aspirin or NSAID in previous 48 hours
- Blood processing
 - Date and time drawn
 - Date and time centrifuged
 - Date and time aliquoted
 - Date and time frozen

Where to start: how to find data items

- Variable search function
- Data collection frequency table
- Participant characteristics over time table
- Data dictionaries for each form
- Form images
- WHI publications

Resources

- Data preparation and use document
- Supplemental documentation on the WHI website
- WHI Help Desk (helpdesk@whi.org)



Department of Health and Human Services

National Institutes of Health

National Heart, Lung, and Blood Institute

WOMEN'S HEALTH INITIATIVE

Thank you to the WHI Participants
for their continued participation!!!