

---

# PLCO DIETARY USER GUIDE

## INTRODUCTION

The Dietary Questionnaire (DQX) and Dietary History Questionnaire (DHQ) are two different instruments given to different groups of participants at different periods of follow-up. Even though many participants in the intervention arm have completed both questionnaires it is strongly recommended not to merge the questionnaire data together. If there is interest in analyzing both questionnaires it should be done separately.

For both questionnaires raw responses were processed into analysis-ready variables in terms of gram intake, pyramid servings, food frequencies per day, etc. See the data dictionaries to review what variables are available and how they were constructed.

The DQX and DHQ datasets are large in size. These files take extra time to download and to work with compared to other PLCO data. Customized datasets of only selected variables can be constructed upon request on a case by case basis.

## PROTOCOL

The Dietary Questionnaire (DQX) was offered to intervention arm participants around the time they were randomized. 82% of the intervention arm completed a DQX. Control arm participants were not offered the DQX. The form may have been completed before (<1%) or after the participant was randomized, but about half (47%) were collected within one month of randomization and most (90%) were collected within three months.

The Diet History Questionnaire (DHQ) was offered to both arms of the trial starting in 1998. 77% of all participants in both arms of the trial completed the DHQ. The form was introduced 5 years into the trial (December 1998). Participants in the control arm randomized prior to December 1998 were offered the DHQ in 1999 or 2000, generally around the anniversary of randomization in either year. Control arm participants randomized in/after December 1998 were offered the DHQ at baseline. Participants in the intervention arm who were randomized prior to December 1995 were offered the DHQ in 1999, generally around the anniversary of randomization. Intervention arm participants randomized in/after December 1995 (third anniversary in/after December 1998) were offered the DHQ generally around their third anniversary of randomization (T3).

For both questionnaires raw responses were processed into analysis-ready variables in terms of gram intake, pyramid servings, food frequencies per day, etc. See the data dictionaries to review what variables are available and how they were constructed.

## COMPLETION

All questionnaires completed are available in the data provided. Completed questionnaires can be identified by DQX\_COMPLETED (> 63,000) and DHQ\_COMPLETED (>118,000). Note, all participants are included in the datasets whether they completed a questionnaire or not.

Most dietary analyses are subset to valid questionnaires only. Four factors are used to identify valid questionnaires:

- A date of completion must be available.
- The date of completion must be prior to the date of death.
- There are less than 8 missing frequency responses.
- Calorie intake is not extreme (top 1% and bottom 1%) for each gender.

Valid questionnaires are identified by the following variables: DQX\_COMPLETEDVALID (>61,000) and DHQ\_COMPLETEDVALID (>113,000).

Table 1: Definition of a Valid DHQ

					<i>Completed Valid DHQ?</i>		
					<i>no</i>	<i>yes</i>	<i>All</i>
<i>All</i>					41441	113446	154887
<i>Completed DHQ?</i>	<i>Missing DHQ Completion Date?</i>	<i>Died Before DHQ Completion?</i>	<i>8 or More Missing/Multiple Frequency Responses on DHQ?</i>	<i>Extreme kCal Consumption on DHQ?</i>			
<i>no</i>	<i>missing</i>	<i>missing</i>	<i>missing</i>	<i>missing</i>	36076	.	36076
<i>yes</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	.	113446	113446
				<i>yes - 1st percentile</i>	911	.	911
				<i>yes - last percentile</i>	1141	.	1141
			<i>yes</i>	<i>no</i>	2967	.	2967
				<i>yes - 1st percentile</i>	273	.	273
				<i>yes - last percentile</i>	45	.	45
	<i>yes</i>	<i>no</i>	<i>no</i>	<i>no</i>	7	.	7
				<i>yes - last percentile</i>	1	.	1
			<i>yes</i>	<i>yes - 1st percentile</i>	2	.	2
	<i>yes</i>	<i>no</i>	<i>no</i>	<i>no</i>	13	.	13
				<i>yes - 1st percentile</i>	1	.	1
				<i>yes</i>	<i>no</i>	4	.

Table 2: Definition of a Valid DQX

					<i>Completed Valid DQX?</i>			
					<i>No</i>	<i>Yes</i>	<i>All</i>	
<i>All</i>					15984	61459	77443	
<i>Completed DQX?</i>	<i>Missing DQX Completion Date?</i>	<i>Died Before DQX Completion?</i>	<i>8 or More Missing Frequency Responses on DQX?</i>	<i>Extreme kCal Consumption?</i>				
<i>No</i>	<i>Missing Data</i>	<i>No</i>	<i>missing</i>	<i>missing</i>	14174	.	14174	
<i>Yes</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>kcal consumption not extreme</i>	.	61459	61459	
				<i>1st percentile</i>	591	.	591	
				<i>last percentile</i>	617	.	617	
			<i>Yes</i>	<i>kcal consumption not extreme</i>	544	.	544	
				<i>1st percentile</i>	41	.	41	
				<i>last percentile</i>	14	.	14	
	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>last percentile</i>	1	.	1
			<i>Yes</i>	<i>Yes</i>	<i>kcal consumption not extreme</i>	2	.	2

## TIME AT ENTRY

The entry dates for the dietary cohorts provided are calculated as the latest of these dates: date of randomization, date completed the baseline questionnaire or the date completed the dietary questionnaire. Note, the order of these events vary which is why the latest is taken. The dietary entry variables are: ENTRYDAYS\_DQX, ENTRYAGE\_DQX, ENTRYDAYS\_DHQ AND ENTRYAGE\_DHQ.

## HEALTH AT QUESTIONNAIRES

Often participants diagnosed with cancer prior to time at entry are excluded. It is presumed that a participant's diet may change if diagnosed with a disease or cancer. Information on co-morbidities and diseases was not asked on the dietary questionnaires,; however, can be found on the baseline questionnaire if diagnosed by then.

Variables identifying participants diagnosed with cancer prior to entry exist for each cohort (PH\_ANY\_DQX and PH\_ANY\_DHQ).