

DPP Data Archive

The full scale DPP data release Archive contains: (1) descriptive documentation; (2) study forms; (3) the data used in the February 2002 *New England Journal of Medicine*¹ study publication (NEJM release), and (4) the full scale DPP data release.

The files in the Archive are organized into the following three directories:

1. Documentation
2. Forms
3. Data

1. Documentation Directory

The Documentation directory contains two subdirectories: *DPP_2006_Data_Release_Docs* and *DPP_2008_Data_Release_Docs*. In addition to the documentation included in the archive, the Manual of Operations is available for approved projects². Also, for a list of study publications, see <http://www.bsc.gwu.edu/dpp/pps/registry.cgi>.

The *DPP_2006_Data_Release_Docs* directory contains documentation from the study and the descriptive documentation developed by and stored in the NIDDK Data Repository. It includes the following documents:

- DPP Data Release Documentation; Data Lock April 1, 2001; Data Release 1, Version 1, January 2006 – Description of the data contained in the NEJM release. (DPP_Data_Release_1_Documentation_2006.pdf). 22 pages.
- DPP Protocol (DPP Protocol.pdf). 119 pages.
- DPP General Description – Description of the clinical trial, deidentification, definition of diabetes, descriptions of variables, and data dictionaries for data releases. (DPP_Study_Website_General_Description.pdf).
- DPP Metadata. (DPP_Study_Website_Metadata.pdf).
- Integrity_FIN_5-10-2006.pdf – Report of an examination of the data conducted by statisticians and quality control specialists at the Repository to provide confidence that the data archive distributed by the Repository is a true copy of the study data. (Integrity_FIN_5-10-2006.pdf). 55 pages.
- Study Timetable – Phase I July 1994 through Study End June 2003. (Study_Timetable.pdf).

¹ The Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med*. 2002 Feb 7;346(6):393-403.

² The DPP DCC requested that the MOOP and Q-forms be released only to approved projects. Although access to these documents can be requested, they are not currently included in the archive at this time, pending NIDDK's resolution of the copyright issues.

The *DPP_2008_Data_Release_Docs* directory contains the following document:

- DPP Data Release Documentation; February 2008 Full Scale Data Release – Description of the data contained in the full scale release. (DPP_Data_Release_Documentation_Full_Scale_2008.pdf). 26 pages.

2. Forms Directory

The Forms directory contains the forms used for data collection. All of the paper forms listed below are in Adobe PDF format. The forms and variables are described in Section 4 of the “Diabetes Prevention Program Data Release Documentation; February 2008 Full Scale Data Release”

[Documentation\DPP_2008_Data_Release_Docs\DPP_Data_Release_Documentation_Full_Scale_2008.pdf]. A listing of the forms and associated datasets is provided in Appendix A. In addition to the forms included in the archive, a set of Questionnaires (Q-forms) is available for approved projects².

Screening Forms

- Form S01 Eligibility Checklist (S01.pdf) – Inclusion and exclusion criteria for potential participants. 6 pages.
- Form S03 Screening Step 2 Inventory (S03.pdf) – BMI, BP, urinalysis, current medications, pregnancy, diabetes, OGTT qualification, OGTT progression, OGTT local results, demographics, complete blood count (CBC) results. 9 pages.
- Form S05 Screening Step 3 Inventory - Start (S05.pdf) – Family history, weight, smoking, aspirin use, cardiovascular and stroke/TIA, other diseases/symptoms, diet, medical history for women, biometrics, BP, dispensing of medication for run-in. 16 pages.
- Form S06 Screening Step 3 Inventory - End (S06.pdf) – Run-in compliance, adverse event assessment, personal and socioeconomic information. 3 pages.
- Form S07 Screening Step 4 Inventory - Randomization (S07.pdf) – Adverse events, pregnancy, prescription medications, final eligibility review, randomization, dispensing of coded medication. 3 pages.

Follow-up Forms (Post-Randomization)

- Form F01 Standard Follow-up Visit Inventory (F01.pdf) – Completed every 3 months between annual follow-up visits. Weight, BP, adverse events, pregnancy, concomitant prescriptions, medication compliance and dispensing, and CBC results. 5 pages.
- Form F02 Major Follow-up Visit Inventory (F02.pdf) – Completed during annual follow-up visits (end-year 1, end-year 2, ...), end of study and at the time of the primary outcome (confirmed development of diabetes): biometrics, BP, adverse events, pregnancy, concomitant prescriptions, medication compliance and dispensing, and CBC results. 6 pages.

- Form F03 Interim Follow-up Visit Inventory (F03.pdf) – Completed at titration visits for coded metformin and follow-up visits when Form F01 and Form F02 are not completed. Reason for visit, adverse events, pregnancy questions, coded medication and BP for hypertension management. 3 pages.
- Form F04 Missed Follow-up Visit Report (F04.pdf) – Missed standard or major scheduled follow-up visit. Date and reason for the missed visit. 1 page.
- Form F05 Medication Adherence Interview (F05.pdf) – Completed when medication adherence is assessed on Form F01 or Form F02, also Form F03 for Month 1. Only for subjects who had taken coded metformin since the previous visit. Data from F05 were transferred to Section H (Coded Medication) of the corresponding Follow-up Visit Inventory. Medication adherence questions. 2 pages.
- Code Sheet for the Medication Adherence Interview (F05_Med_Adhere_Codes.pdf) – Reasons for noncompliance, strategies for improved compliance. 2 pages.
- Form F06 Home Visit Inventory (F06.pdf) – Inactive participants off their coded medication for any Mid-Year or Annual visit conducted outside the DPP clinic. Adverse events, prescription medications. 2 pages.

Lifestyle Intervention Group Forms

- Form L03 - Lifestyle Contact - In Person (L03.pdf) – Completed for all in-person contacts: Includes nature of session (type and duration), self-monitoring information, physical activity and weight status. 5 pages.
- Form L04 - Lifestyle Physical Activity Log (L04.pdf) – Completed for each supervised physical activity session. Type of activity, list of attendees. 1 page.
- Form L05 - Lifestyle Group Session Log (L05.pdf) – Completed for each group session. Type of activity, list of attendees. 1 page.
- Lifestyle intervention codes (L03_L04_L05_Lifestyle_Codes.pdf). Exercise and lifestyle codes. 13 pages.

Event Data forms

- Form E01 Adverse Event Report (E01.pdf) – New symptoms, injuries, illness, side effects, or worsening of pre-existing conditions. Adverse event classification and possible interventions (metformin, diet, exercise). 4 pages.
- Form E02 Serious Adverse Event Report (E02.pdf) – Hospitalization, cancer, disability, death, use of metformin, treatment, use of insulin, outcome. 4 pages.
- Form E04 Pregnancy Confirmation Report (E04.pdf) – For subjects who had a positive pregnancy test. Metformin discontinued. 2 pages.
- Form E05 Pregnancy Outcome Report (E05.pdf) – For subjects whose pregnancy had ended. 2 pages.
- Form E06 Mortality Event Report (E06.pdf) – Circumstances of death, use of metformin, causes of death. 3 pages.

Heart Disease Risk Form

- Form R04 CHD Risk Status Report (R04.pdf) – Coronary heart disease risk, risk

factors, need to unmask lipid profile. 2 pages.

Troglitazone Forms

- Form TR1 Participants Randomized to Troglitazone Follow-Up Visit Inventory (TR1.pdf) – Mid-year and annual follow-up visits for participants randomized to troglitazone. (End-month 6, 12, 18, ...). Weight, BP, adverse events and concomitant medications. 3 pages.
- Form TR02 Participants Randomized to Troglitazone Group Session Log (TR2.pdf) – Completed for each troglitazone group session. Type of activity, list of attendees. 1 page.

3. Data Directory

The Data directory contains two subdirectories: *DPP_Data_NEJM_2002* and *DPP_Data_2008*.

The *DPP_Data_NEJM_2002* subdirectory contains the datasets used in the NEJM publication¹, which was based on participant visits in DPP as of April 1, 2001. The data are contained in the following three SAS v7 format datasets:

- *dpp_b1v1*: One record for each of the 3154 randomized participants. Each record includes baseline variables and event variables requiring only one value such as diabetes status.
- *dpp_f1v1*: One record per participant for each semi-annual visit attended. Each record includes followup variables that were measured sequentially over time.
- *dpp_h1v1*: One record for each hospitalization for each participant who had at least 1 hospitalization.

The values for categorical variables in the datasets are provided in “DPP formats.doc”.

See “Diabetes Prevention Program Data Release Documentation Data Lock: April 1, 2001” [Documentation\DPP_2006_Data_Release_Docs\DPP_Data_Release_1_Documentation_2006.pdf] for additional information on the structure and contents of these datasets and information on statistical considerations of the analysis.

The *DPP_Data_2008* subdirectory contains the full-scale data. The data are contained in the following two subdirectories:

- **Form Data**: Data collected on forms, as specified in section 4.4-4.10 of “Diabetes Prevention Program Data Release Documentation February 2008 Full Scale Data Release” [Documentation\DPP_2008_Data_Release_Docs\DPP_Data_Release_Documentation_Full_Scale_2008.pdf]. Contains a **Data** folder and a **Data Dictionary** folder. The **Data** folder contains the .XPT files listed below. There is an .XPT file for each of the 30 Forms; it contains the data collected using that form. The **Data Dictionary** folder contains the .doc files listed below. There is a .doc file for each form; it describes the variables in

the form. Instructions for opening the .XPT file can be found on page 5 of the “Diabetes Prevention Program Data Release Documentation February 2008 Full Scale Data Release” document.

•

	<i>Data</i>	<i>Data Dictionary</i>
1.	E04.XPT	E04.DOC
2.	E05.XPT	E05.DOC
3.	F01.XPT	F01.DOC
4.	F02.XPT	F02.DOC
5.	F03.XPT	F03.DOC
6.	F04.XPT	F04.DOC
7.	F05.XPT	F05.DOC
8.	F06.XPT	F06.DOC
9.	L03.XPT	L03.DOC
10.	L04.XPT	L04.DOC
11.	L05.XPT	L05.DOC
12.	Q01.XPT	Q01.DOC
13.	Q02.XPT	Q02.DOC
14.	Q03.XPT	Q03.DOC
15.	Q04.XPT	Q04.DOC
16.	Q05.XPT	Q05.DOC
17.	Q06.XPT	Q06.DOC
18.	Q08.XPT	Q08.DOC
19.	Q09.XPT	Q09.DOC
20.	Q10.XPT	Q10.DOC
21.	Q12.XPT	Q12.DOC
22.	Q13.XPT	Q13.DOC
23.	R04.XPT	R04.DOC
24.	S01.XPT	S01.DOC
25.	S03.XPT	S03.DOC
26.	S05.XPT	S05.DOC
27.	S06.XPT	S06.DOC
28.	S07.XPT	S07.DOC
29.	TR1.XPT	TR1.DOC
30.	TR2.XPT	TR2.DOC

- **Non-Form Data:** Data not collected on forms but for which datasets are included in this release, as specified in section 4.11 – 4.12 of “Diabetes Prevention Program Data Release Documentation February 2008 Full Scale Data Release” [Documentation\DPP_2008_Data_Release_Docs\DPP_Data_Release_Documentation_Full_Scale_2008.pdf]. Contains a *Data* folder and a *Data Dictionary* folder. The *Data* folder contains the .XPT files listed below. The *Data Dictionary* folder contains the .doc files listed below. There is a .doc file for each dataset; it describes the dataset variables.

For additional information on the structure and contents of these datasets and statistical considerations of the analysis, see the “Diabetes Prevention Program Data Release

Documentation February 2008 Full Scale Data Release” document. Also, instructions for opening the .XPT files can be found on page 5 of this document.

Data

1. BASEDATA.XPT
2. CT.XPT
3. EVENTS.XPT
4. LAB.XPT
5. NCC.XPT
6. QWB.XPT

Data Dictionary

- BASEDATA.DOC
- CT.DOC
- EVENTS.DOC
- LAB.DOC
- NCC.DOC
- QWB.DOC

Appendix A: Listing of DPP Forms and Datasets

DPP Forms and Data			
Form Name	Data File Name	Variables	Size
E01 Adverse Event Report	---	---	---
E02 Serious Adverse Event Report	---	---	---
E04 Pregnancy Confirmation	E04.XPT	11	6 KB
E05 Pregnancy Outcome	E05.XPT	21	6 KB
E06 Mortality Event	---	---	---
F01 Standard Follow-up Visit Inventory	F01.XPT	43	4.86 MB
F02 Major Follow-up Visit Inventory	F02.XPT	72	2.13 B
F03 Interim Follow-up Visit Inventory	F03.XPT	32	1.4 MB
F04 Missed Follow-up Visit Report	F04.XPT	9	111 KB
F05 Medication Adherence Interview	F05.XPT	19	2.35 MB
F06 Home Visit Inventory	F06.XPT	17	3.0 KB
L03 Lifestyle Contact - In Person	L03.XPT	62	10.9 MB
L04 Lifestyle Physical Activity Log	L04.XPT	33	924 KB
L05 Lifestyle Group Session Log	L05.XPT	33	193 KB
Form Q01 Beck Questionnaires	Q01.XPT	50	1.12 MB
Form Q02 Health Survey Questionnaire	Q02.XPT	42	1.65 MB
Form Q03 Modifiable Activity Questionnaire	Q03.XPT	222	6.53 MB
Form Q04 Low Level Physical Activity Recal	Q04.XPT	50	3.04 MB
Form Q05 Nhanes III Physical Activity Scale	Q05.XPT	47	523 KB
Form Q06 Retention And Treatment Monitoring Measures	Q06.XPT	71	2.64 MB
Form Q08 Interval History Questionnaire	Q08.XPT	61	1.70 MB
Form Q09 DPP-Specific Support Measure - Baseline Visit	Q09.XPT	7	224 KB
Form Q10 DPP-Specific Support Measure - Follow-up Visits	Q10.XPT	12	614 KB
Form Q12 Economic Evaluation Questionnaire	Q12.XPT	51	441 KB
Form Q13 Urinary Incontinence Questionnaire	Q13.XPT	16	249 KB
Form R04 CHD Risk Status Report	R04.XPT	15	1.37 MB
Form S01 Eligibility Checklist	S01.XPT	34	723 KB
Form S03 Screening Step 2 Inventory	S03.XPT	56	754 KB
Form S05 Screening Step 3 Inventory - Start	S05.XPT	173	1.58 MB
Form S06 Screening Step 3 Inventory - End	S06.XPT	23	324 KB
Form S07 - Screening Step 4 Inventory - Randomization	S07.XPT	28	401 KB
Does Not Apply	TR1.XPT	25	383 KB
Does Not Apply	TR2.XPT	33	27 KB
Does Not Apply	BASEDATA.XPT	6	269 KB
Does Not Apply	CT.XPT	7	164 KB
Does Not Apply	EVENTS.XPT	11	332 KB
Does Not Apply	LAB.XPT	33	5.14 MB
Does Not Apply	NCC.XPT	150	6.80 MB
Does Not Apply	QWB.XPT	9	553 KB