

# DPM Sampling, Study Design, and Calculation Methods

## Table of Contents

<i>DPM Sampling, Study Design, and Calculation Methods</i> .....	1
<i>Facility Sample Frame</i> .....	1
<i>Facility Sampling Strata</i> .....	2
<i>Patient Sampling</i> .....	3
<i>Sampling Weights</i> .....	3
<i>DOPPS Data Elements</i> .....	4
<i>How Data are Reported in the DPM</i> .....	4
<i>Evaluating the Representativeness of the DPM sample</i> .....	5
<i>Data Entry and Software</i> .....	6
<i>References</i> .....	6

The DOPPS Practice Monitor aims to provide nationally representative estimates using a complex sample design. An updated methods paper, providing details of the DPM study design, sampling approach, and facility sample characteristics has been published recently<sup>1</sup>. This paper also includes information demonstrating that the sample of US facilities in the DPM yields results representative nationally and by facility type (dialysis organization size, rural/urban, free standing/hospital based), achieving coverage similar to the CMS sample frame at average values and tails of the distributions for key measures and patient characteristics. Additional comparisons, showing close correspondence of DPM results with national ELab statistics for late 2010, are provided near the end of this document.

Although great efforts are made to provide nationally representative statistics in the DPM, it is important to note that sampling variability, the use of survey weights, and calculation methods may produce results that differ slightly from other nationally representative or registry data. Selected aspects of the DPM study design and calculation methods are summarized below for readers of the DPM.

### ***Facility Sample Frame***

In each DOPPS country, facilities are randomly selected from a list of all dialysis facilities (i.e., sample frame) treating at least 20 adult in-center hemodialysis patients ( $\geq 18$  years of age). This facility sample frame was based on the 2008 Dialysis Facility Compare database (DFC; REF:

www3.cms.gov/DialysisFacilityCompare), created from CMS data by Arbor Research Collaborative for Health and the University of Michigan Kidney Epidemiology Cost Center. This DFC database provides demographic and quality-of-care measures for all Medicare-certified dialysis facilities in the United States collected through the end of 2007. Data from the DFC were augmented with facility patient counts obtained from the July 2008 Annual Facility Survey database (AFS; REF: Form CMS-2744). A data use agreement with CMS provided access to CMS data.

Of 4,959 facilities listed in the DFC database, 4,343 (88%) served at least 20 chronic adult ( $\geq 18$  years old) in-center hemodialysis patients and comprised the final DPM sample frame of facilities eligible for study participation. The restriction to facilities treating at least 20 patients, due to resource limitations, is readily justifiable because these facilities served 97% of U.S. hemodialysis patients. From this sample frame, approximately 145 dialysis units were targeted for DPM recruitment to meet the study's scientific goals.

### ***Facility Sampling Strata***

Random facility selection was performed within pre-specified strata of different HD facility types. Prior DOPPS experience has demonstrated that data from 20 randomly selected facilities yield reliable estimates to accurately represent national hemodialysis practices. Detailed DOPPS methods have been published previously<sup>2,3</sup>. Based on this experience, targets were set to select a minimum of 20 facilities per category in the following three strata:

**(A) DIALYSIS ORGANIZATION SIZE** – Based on the number of units comprising a dialysis organization, dialysis units were classified as part of an LDO (large dialysis organization; >1,000 dialysis facilities), MDO (medium-size dialysis organization; 10-<1,000 facilities), and independent or SDO (small-size dialysis organization; <10 dialysis facilities). Notably, the LDO stratum is composed of only the two largest U.S. dialysis organizations; currently, there are no U.S. dialysis organizations comprised of 300-1000 dialysis facilities. Also, we are unaware of a standard cut-point to distinguish SDO from MDO; the cut-point of 10 facilities was used to meet the needs of the DPM sampling design.

**(B) HOSPITAL-based / NON-HOSPITAL-based** – All dialysis units were categorized as either hospital-based or free-standing, non-hospital based dialysis units. Note that all hospital-based units in CMS data were SDO/independent facilities, based on the definition above.

**(C) RURAL / NON-RURAL LOCATION** – Dialysis units were grouped into 3 categories based on the 2004 census tract-based rural-urban commuting area (RUCA) codes available from WWAMI Rural Health Research Center at the University of Washington, Seattle WA

(<http://depts.washington.edu/uwruca/index.php>). Zip codes were used to assign facilities to a particular RUCA code. Facilities were grouped into the 3 categories of: (i) Remote Rural, defined as having RUCA code 7.0-10.6 (population < 10,000) with >45 minutes travel time to the nearest urbanized area of >50,000 population), (ii) Non-remote Rural, defined as having RUCA code 7.0-10.6 (population < 10,000) with <45 minutes travel time to the nearest urbanized area of >50,000 population), and (iii) Non-rural, defined as having a RUCA code <7.0 (i.e., population  $\geq$  10,000). Travel time of less than versus greater than 45 minutes to the nearest urbanized area of >50,000 population was taken from this same University of Washington data source.

### ***Patient Sampling***

To reduce the data collection burden at recruited study sites, data are collected from 20-40 randomly selected chronic in-center HD patients at each participating study site, using a selection algorithm that selects a larger number of patients with increasing size of the dialysis unit to yield an overall average of around 30 study patients per facility. Patients departing the study (e.g., due to death, transfer to other facilities, switch to peritoneal dialysis or home HD, transplantation, or recovery of renal function) are replaced at four month intervals.

Replacements are determined by random selection of patients new to the facility since the prior selection, providing prevalent cross-sections that are representative of patients in each facility at every four month interval. Informed patient consent is obtained in accordance with local requirements.

### ***Sampling Weights***

Facility sampling weights, patient sampling weights, and post-stratification weights are applied to report data that are representative at the national and facility stratum levels. To obtain  $\geq$ 20 facilities in each stratum category, over-sampling of facilities occurs in some categories. The fraction of facilities within each stratum that are participating in the DOPPS is termed the facility sampling fraction. The reciprocal of the facility sampling fraction serves as the facility sampling weight. Within each facility, data are collected on a random sample of 20-40 chronic adult patients. The fraction of all of the facility's chronic adult HD patients that are participating in the study is termed the patient sampling fraction. The reciprocal of the patient sampling fraction serves as the patient sampling weight. Additionally, post-stratification weights based on multiple regression are computed for each patient to correct for disproportionate consent rates among basic demographic variables collected on all facility patients.

### ***DOPPS Data Elements***

At study entry, patient data are collected which include demographics, detailed comorbidities, vascular access, medication use/dose, and laboratory values. Patient-level data are updated every four months including: laboratory values (recorded once monthly); medication use/dose (current medication list, as well as monthly use/dose for mineral metabolism and anemia-related medications); vascular access procedures, other clinical procedures, and hospitalizations; dates and cause of death; and reasons for study departure. Patient self-reported data (e.g., quality of life, depression symptoms, satisfaction with care, and others) are collected yearly. Facility survey data related to practice preferences, services offered, and staffing are collected yearly from the facility medical director and nurse study manager. DOPPS data collection instruments are updated annually.

### ***How Data are Reported in the DPM***

Demographic and comorbidity data will be reported annually. Vascular access use will be reported every four months. Anemia and mineral bone disorder medication use will be reported monthly as a reflection of medication use either in the given study month and/or during the prior three months. Use of several classes of cardiovascular and anti-hypertension agents will be reported once every four months. Doses will be reported for erythropoiesis-stimulating agents and intravenous iron. Routinely measured laboratory data will be reported for each month and as the mean of the last 2 to 3 monthly values. Less frequently measured laboratory data (e.g., PTH, ferritin) will be reported as the most recent monthly value and the mean of the last 1 to 3 monthly values. The proportion of patients with serum ferritin and parathyroid hormone measurements will also be reported for one-month and three-month periods.

Continuous variables (e.g., hemoglobin levels) will be presented as simplified box plots showing the median (50<sup>th</sup> percentile) value as a large dot, with auxiliary lines indicating the 10<sup>th</sup>-25<sup>th</sup> and 75<sup>th</sup>-90<sup>th</sup> percentiles. Categorical variables will be presented as column (two levels) or stacked column charts with clinically meaningful cut-points. Annotated examples of these types of charts are available on the DPM website: <http://www.dopps.org/DPM/Understanding.pdf>

Annualized rates for clinical outcomes (mortality, hospitalizations, transfers out of the dialysis unit, transplant, and withdrawal from dialysis) were calculated as the number of events over a 30 day period starting on the 15th of the listed month divided by the time at risk among patients at risk of an event. Note that as the DOPPS uses a random sample of dialysis facilities treating at least 20 in-center hemodialysis patients, the rates presented may differ somewhat from other published sources.

Emerging trends will be monitored for variables reported on the DPM website. Trends will be analyzed by several approaches, including longitudinal analyses. Tests of location (T tests, quantile regression) that account for repeated measures will be used to assess significant differences from baseline (e.g., 2010) values. After a sufficient number of follow-up periods have been obtained, additional modeling techniques including spline regression will be used to describe possible non-linear changes. Trends will be monitored at a national level, as well as within and between facility strata. A summary document provided with each DPM update will highlight notable findings.

Certain clinical data will not be provided on the website, but will form the basis for multivariable analyses and peer-reviewed publications. These include rates of clinical procedures and other events, facility survey data, and selected patient self-reported data.

### ***Evaluating the Representativeness of the DPM sample***

To assess the extent to which the sample of DPM facilities is representative of U.S. HD facilities overall, analyses have been performed comparing 2007 CMS data for 137 DPM facilities participating in the DOPPS in September, 2010 with the 2007 CMS data for all U.S. facilities treating at least 20 in-center chronic HD patients (n=4,343 facilities)<sup>1</sup>. Analyses were limited to chronic HD patients  $\geq 18$  years of age. Comparisons were performed for 8 patient characteristics or laboratory measures available in the CMS data. These items were chosen because they are part of the 2012 Quality Incentive Program, were identified as important in a recent GAO Report {cite}, or are otherwise expected to be potentially affected by the Prospective Payment System. Results of these analyses indicate the DPM is nationally representative of U.S. HD patients with nearly all of these tested measures, except for being under-represented in the percent of black HD patients (26% black patients in DPM vs. 36% in CMS data, overall). Although under-represented in overall percent of black HD patients, the DPM facility sample provides ample representation of the entire distribution of the facility percent of black patients and thus will be able to demonstrate changes or effects related to the percent of black patients within U.S. HD facilities. Additional details and results regarding these analyses are described in Robinson, et al<sup>1</sup>.

In June 2011, ESRD Network 11 Elab Project released their "National 2010 and Trends" report<sup>4</sup>. The Elab project collects 4<sup>th</sup> quarter 2010 laboratory measurements from approximately 387,000 (97% of the US dialysis population) patients and therefore provides nearly complete national data against which DPM estimates from the baseline 2010 period can be validated. Elab reports averages based on the mean of the first laboratory value in October, November, and December 2010. The mean of up to three monthly values from October-December 2010 is

also reported in the DPM (as "Dec 2010", where available), providing a similar methodology and enabling a fairer comparison. The following table represents the comparison of Elab and DPM estimates for demographic characteristics and laboratory variables in common.

Patient Characteristics	Elab 2010	DPM Dec 2010
Age, 18-54 yrs, (%)	30.2	29.2
Age, 55-74 yrs, (%)	48.1	47.9
Age, ≥75 yrs, (%)	21.7	23.0
Male, (%)	55.4	55.4
Black, (%)	37.6	30.8
Diabetic, (%)	54.6	62.1 <sup>a</sup>

<sup>a</sup>Ascertainment of diabetes status differs between Elab and DPM; DPM incorporates use of diabetes-related medications as part of ascertainment in addition to indication of diabetes in medical record.

Laboratory Measures	Elab 2010	DPM Dec 2010
Mean Hemoglobin (Hgb), g/dL	11.5	11.5
Hgb <10 g/dL, (%)	6.6	5.5
Hgb 10-12 g/dL, (%)	68.4	71.1
Hgb >12 g/dL, (%)	25.0	23.3
TSAT ≥20%, (%)	87.0	87.6
URR ≥ 65%, (%)	91.1	92.6
S. Albumin ≥4.0 g/dL, (%)	39.1	40.3
S. Phosphorus, 3.5-5.5 mg/dL, (%)	55.3	58.0
S. Calcium, 8.4-9.5 mg/dL, (%)	64.1	64.9

Values shown are three-month averages for each measure.

### Data Entry and Software

Anonymous raw data are collected from DOPPS study sites via either (1) electronic data transfer from databases, (2) data entry into a proprietary web-based application (DOPPSLink), or (3) paper forms keypunched in-house. Range and consistency checks are performed on all data. Statistical analyses are performed using SAS 9.2 (SAS Institute, Cary, NC).

### References

- [1] Robinson BM, Fuller DS, Dykstra DM, Albert JM, Gillespie BW, Tentori F, Turenne M, Port FK, Pisoni RL. (2011). "The DOPPS Practice Monitor: Rationale and Methods for an Initiative to Monitor the New US Bundled Dialysis Payment System." *Am J Kidney Dis*, 57, pp. 822-831
- [2] Young EW, Goodkin DA, Mapes DL, Port FK, Keen ML, Chen K, Maroni BL, Wolfe RA, and Held PJ. (2000). "The Dialysis Outcomes and Practice Patterns Study (DOPPS): An international hemodialysis study." *Kidney Int*, 57 Suppl. 74s, pp. S74-S81.

[3] Pisoni RL, Gillespie BW, Dickinson DM, Chen K, Kutner MH, and Wolfe RA. (2004). "The Dialysis Outcomes and Practice Patterns Study (DOPPS): Design, data elements, and methodology." *Am J Kidney Dis*, 44, pp. 7-15.

[4] ESRD Network 11 (2011). "National 2010 and Trends Elab Report".  
[http://www.esrdnet11.org/Elab/elab\\_national\\_2010\\_and\\_trends\\_report.pdf](http://www.esrdnet11.org/Elab/elab_national_2010_and_trends_report.pdf); retrieved Nov 8, 2011.