

# **Time Study Proposal**

## **School Based Medicaid Administrative Match Program Washington State**

**October 2003**

**Designed and Prepared by:  
*Washington State Department of Social and Health Services*  
Management Services Administration  
Research and Data Analysis Division  
Dario Longhi, Ph.D., Senior Research Manager**



# Table of Contents

<b>TIME STUDY PROPOSAL</b>	<b>Pages</b>
<b>Overview of Proposal</b> .....	<b>1</b>
Study Design .....	1
Sampling of Staff .....	1
Sampling of Days .....	1
Study Process .....	3
<b>The Modified RMS Design</b> .....	<b>5</b>
Random Moment Sampling (RMS) .....	5
Modifications .....	5
The Statistical Sampling Frame for Selecting Staff .....	7
Sampling Decisions Produce Conservative Statewide Estimates .....	10
Major RMS features are Preserved to Provide Unbiased Estimates .....	10
<b>Detailed Features of the Proposed Design</b> .....	<b>13</b>
The Sampling Universes of Eligible Persons and Time are Clearly Defined and Verifiable .....	13
Calculation of Overall Statewide Estimates Involve Weighted Averages Due to Different Sampling Fractions and Sizes of Strata .....	13
Allocation of Funds to Individual Large and Very Large School is Based on Two Time Estimates: a district specific one for “Designated” Staff and a pooled one for “Teachers and Other Staff.” .....	13
Sample Sizes in the Modified Design Compared to Those in a Simple Random Sample RMS Design .....	14
Summary of Reasons for Modifying the Simple Random Sample Design .....	15
<b>References</b> .....	<b>17</b>



## Overview of Proposal

The purpose of this time study is to measure, quarterly, the amount of time spent by school staff on Medicaid reimbursable activities. A modified Random Moment Sampling (RMS) design is proposed, where both staff and time are randomly sampled, and the frequency of certain activities is measured. Staff samples provide time estimates based on 5% error and 95% confidence, with a conservative margin. This RMS design makes conservative assumptions and aims to satisfy desired levels of statistical validity.

Since Medicaid reimbursable activities represent a very small proportion of most school staff time – about 1-4 % of the total work day, purely random time moments would capture inapplicable activities 96-99% of the time. Surveying staff at any random moment concerning their activities would be quite distracting for most respondents (while teachers are teaching, counselors are counseling, etc.) and a very inefficient way of collecting such information.

**Study design** - A modification of the RMS design is proposed whereby:

- Eligible school staff are identified by OSPI program and funding codes (Attachment 2);
- School staff are randomly selected each quarter for each school district (Attachment 3);
- School days are randomly selected each quarter, independently for each school district;
- A log of activities is entered on a standard form by selected staff on each selected day.

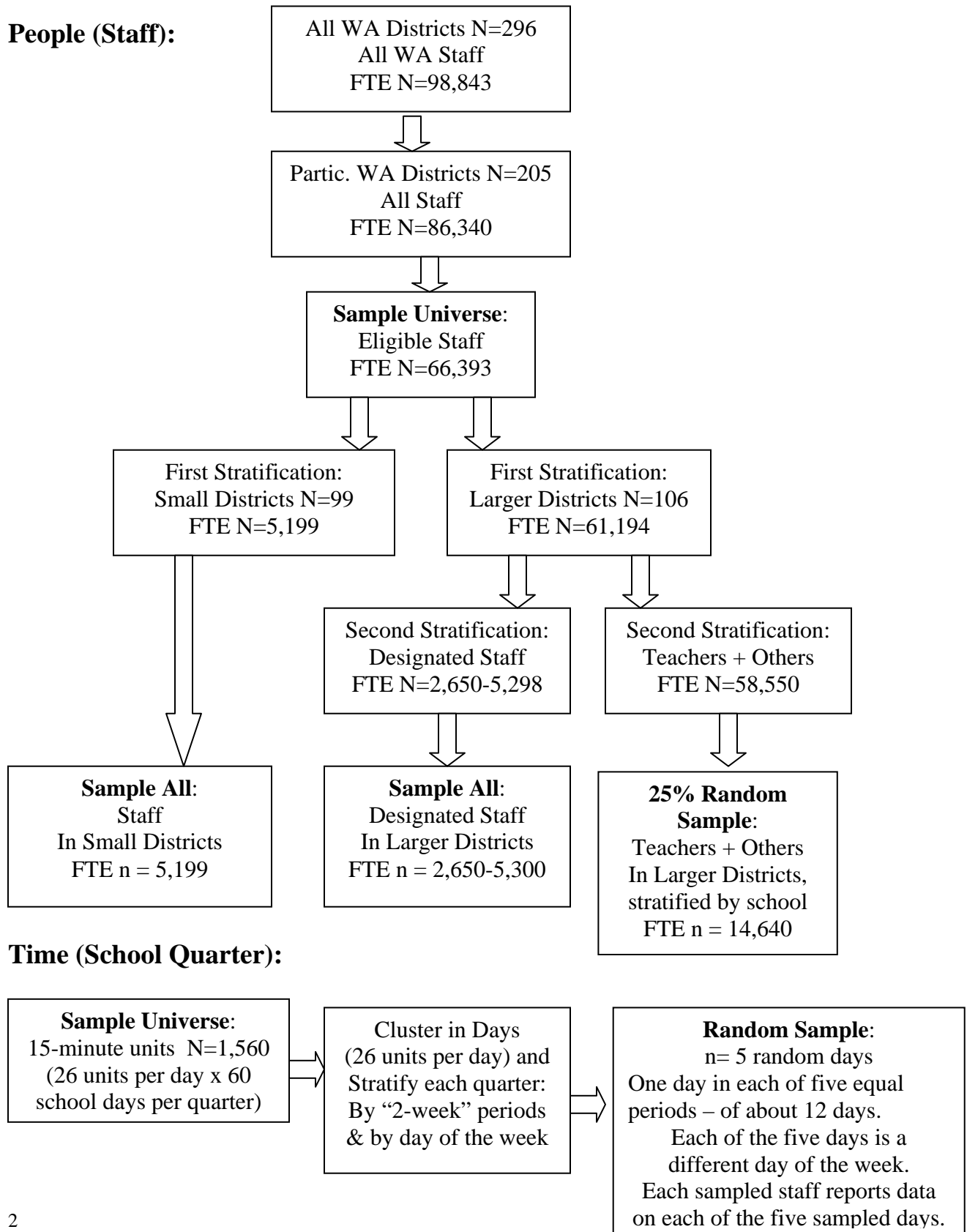
This modified design still insures random, unbiased, measurement, but in a less intrusive, more efficient, way than a completely random moment sampling design. A minority of staff is affected on a few randomly selected days each quarter. The sampling universe of appropriate “school days” and “eligible” staff is checked each quarter. Days and staff are selected at random by a computerized system operated by an agency external to the school districts: the Department of Social and Health Services (DSHS).

**Sampling of staff** - Further efficiencies are gained, and more accurate time estimates are obtained, by stratifying the population of staff into three categories (see Figure 1) and selecting:

- All eligible staff in the 99 smaller school districts -- a statewide sample size of 5,199 with an average of approximately 53 FTE per district;
- All ‘designated’ staff in the 106 larger school districts – representing a range of 4-8% of all eligible staff who do most of the outreach, and actual connecting of students to Medicaid paid services (averaging 25-50 FTE per district = 2,650-5,300 statewide); and
- 25% “teachers and other support staff” -- selected at random in the 106 larger school districts -- who do mainly referral and monitoring activities for ‘their’ students -- a statewide sample size of 14,640 with an average of approximately 138 FTE per district.

**Sampling of days** - It is proposed that DSHS select a random sample of five school days specific to each school district prior to each quarter from each school district’s calendar. A computer program has been written to select at random five school days distributed across five “two week” periods in the quarter and represent every day of the week (Monday through Friday) every quarter (See Attachment 3). Each staff sampled in each district will provide five days of data every quarter.

**Figure 1**  
**Washington State Modified RMS Design**



**Study process** - To conduct this time study, a series of steps needs to be followed each quarter. Some steps are done by school districts, some by DSHS (See Figure 2; Attachment 7, Contracts).

**At the beginning of each quarter:**

- |                  |  |
|------------------|--|
| School Districts | Verify their calendars of “school days;”<br>List all eligible school staff by school, divided into “designated” and “teacher and other staff” for larger districts, and send the lists to DSHS;<br>Train all eligible staff (Attachments 4 and 5). |
| DSHS             | Selects samples of days and staff, and sends them back to district coordinators (Attachment 3);<br>Monitors training (Attachment 6);<br>Supplies data collection forms (Attachment 1).   |

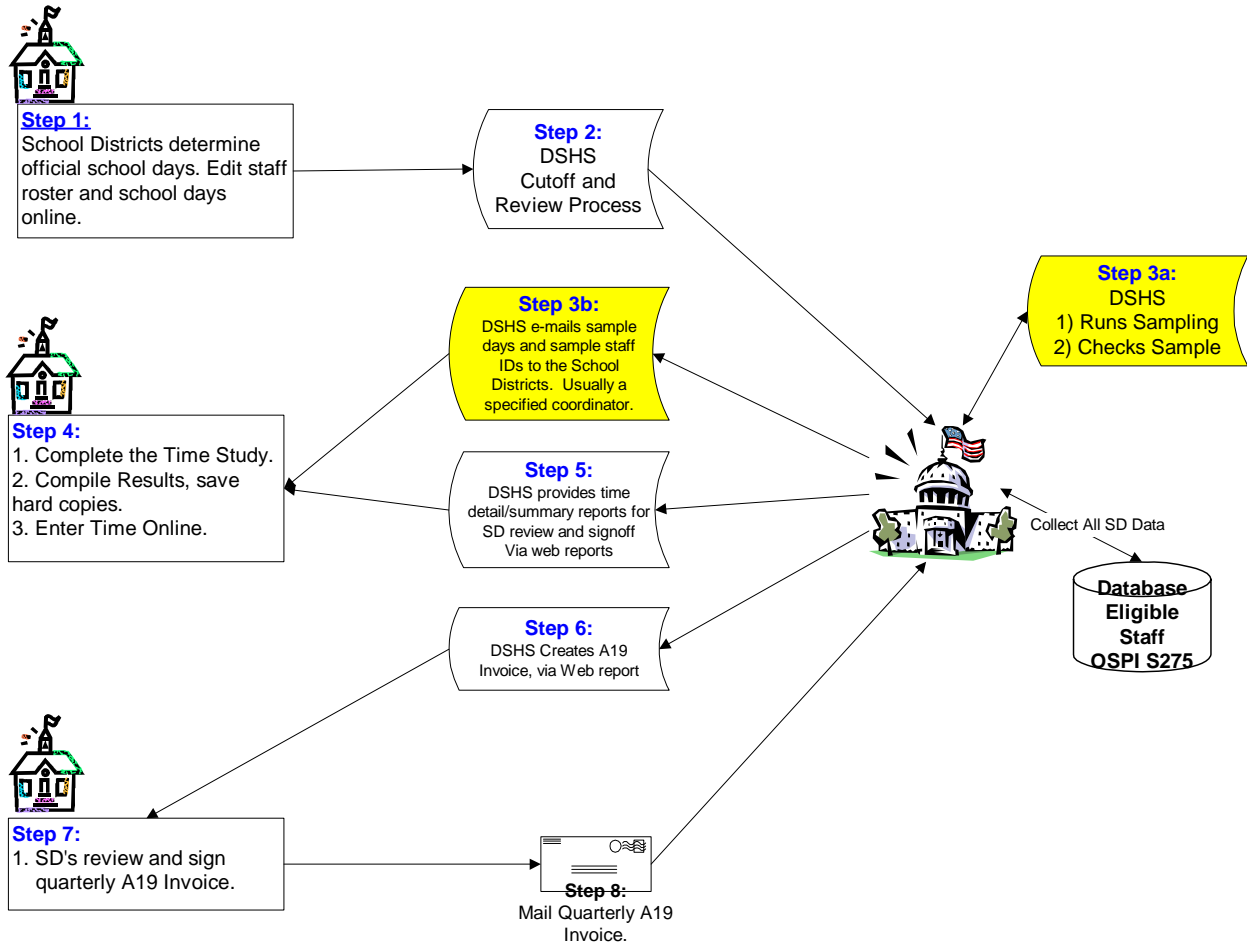
**During each quarter:**

- |                  |  |
|------------------|--|
| School Districts | Two to five days before each sampled day, district coordinators send notification and forms to sampled staff;<br>Collect forms from staff within two days after sampled day;<br>Send raw data electronically to DSHS (web based form and data entry system are being developed). |
| DSHS             | Monitors timeliness of returns (Attachment 6);<br>Summarizes results and calculates average time spent (non-returned forms are assumed to represent “0” time spent);<br>Returns individual results for each day electronically to School Districts.                              |

**At the end of each quarter:**

- |                  |  |
|------------------|--|
| DSHS             | Documents compliance with design and sampling methodology (see Monitoring, Attachment 6);<br>Calculates weighted pooled average results for larger school districts (Attachment 9) and individual results for smaller school districts;<br>Calculates statewide weighted time estimates and precision;<br>Sends quarterly results and draft invoices to district coordinators.   |
| School Districts | Check, sign and submit invoices to DSHS. The invoices will be based on:<br><i>For 99 small districts:</i> The time measurement for each of the 99 districts;<br><i>For 106 large districts:</i> The time measurement for designated staff in each district,<br>plus<br>The time measurement for teachers and other staff in each district -- estimated by multiplying individually measured “designated” time for each district by the ratio of pooled “teacher” time to pooled “designated” time (see example in Attachment 8). |

**Figure 2  
Quarterly Process**



**Assumptions:**

1. Annual Procurement of District school calendars.
2. Schools maintain employee rosters using DSHS templates.
3. Step 2 is not under HIPAA.

**Process Characteristics:**

1. DSHS determines sample days and sample staff.
2. DSHS Collects time data via online entry screen.
3. DSHS provides A19 for each school district.



# The Modified RMS Design

## Random Moment Sampling (RMS)

The proposed time study is a modified version of Random Moment Sampling (RMS) which is the most efficient tool for periodically estimating time spent on various activities by persons conducting routine jobs (See Bolstein, 1986; Niebal, 1972).

RMS uses a well-defined set of activities into which work can easily be divided. This study proposes to identify Medicaid-related activities and collect data on them through a standard data collection form. (see Attachment 1: Data Collection Form and Activity Codes)

RMS defines the sample universe, the population, as the product of people and time: the product of all persons and all possible time moments. It then samples a set of random person/moments (a particular time for a particular person) and infers the sample results to the population.

## Modifications

The proposed design is a modified version of the RMS general design. Although not a completely random design for sampling moments, it is a random sampling of staff and days within strata. Stratifying the sample universe of persons and days improves the precision of estimates compared to simple random sampling.

*For persons*, the model proposes:

1. A stratification of school districts by size – two main strata composed of school staff working in smaller versus larger districts (less than/ more than 150 FTE) and a further subdivision among the larger districts into a large and very large category (less than/ more than 1,550 FTE). See Attachment 9 for a frequency distribution table of school districts by size and of time spent on Medicaid Activities in 2001-2002 across school districts of different sizes. Note the major differences in average time spent by staff in small districts (6.2%), large districts (4.7%) and very large districts (1.9%).
2. A stratification of staff within larger school districts, first by school, and then, within each school, by duties and tasks that they perform – two strata per school composed of a category of ‘designated’ staff (performing a variety of specialized tasks linking students to Medicaid eligibility and services), and a second category composed mainly of teachers and some support staff (performing mainly referral and monitoring tasks). Past time studies indicate that the number of ‘designated’ staff is small, about 4-8% overall, and that they spend more time, on average, on Medicaid activities than teachers and other support staff (about 10-20% versus 1-4%, respectively).

**For time**, the model proposes clustering time moments into full “school days,” since only a very small part of each day is spent on Medicaid activities (about 3% on average) and sampling days is more practical:

1. 15-minute intervals will be sampled clustered into five full days generating a total of 130 time moments. The days will be randomly selected for each school district based on the district school calendars (See Attachment 3).
2. Each sampled staff member in a participating school district will complete full-day time logs for all the five randomly selected days for that district, recorded in 15-minute intervals.
3. Medicaid activities, consistent with CMS guidelines, will be clearly defined. (See the proposed form and activities descriptions in Attachment 1, Documentation in Attachment 4, Training in Attachment 5 and DSHS contracts with School Districts in Attachment 7.)

The model also proposes to sample days with stratification. Some exploratory work with expert informers suggests that time spent varies most by

1. early/late period of the quarter – the ‘first two weeks’ versus the ‘last two weeks’ versus the ‘middle’ period of the quarter, and
2. day of the week – referral activities are said to happen earlier in the week, time consuming connecting activities later in the week.

Therefore, the model proposes a stratification of days that provides a proportional representation of all time periods in the quarter and of all the days of the week:

1. A stratification of five equal “12-day” periods throughout the school quarter – to evenly distribute the sampled school days (avoiding possible grouping of simple randomly selected days at the beginning, middle or end of the quarter).
2. A stratification of days of the week (Monday to Friday) – to guarantee that all the days of the week are included in the sample. (See the description of the automated random sampling procedures for selecting days in the quarter in Attachment 3).

First quarter results using this approach will provide estimates of both the average time spent each day of the quarter and also the variance in time spent across the 60 days in the quarter. For example, sample size for each day in the quarter will be approximately 1,200 teachers (14,400/60). The sample sizes will be sufficient to provide fairly accurate estimates of time spent by period of the quarter and by day of the week (n=14,400 teachers aggregated for each period and each day of the week).

**Proviso:** Future studies on first quarter results will test the assumptions of independence of people and time and may recommend modifications in sampling calculations and consequent sample sizes. These modifications may include different 5 days per staff member within each school district based on a random selection of days. Depending on first quarter results this may become necessary for larger school districts. For these districts it may be practical to sample different five day periods for each type of school: for all elementary schools, middle schools, high schools and alternative schools.

*Statistical note:* The formula used to establish sample sizes assumes a “yes/no” dichotomous variable to designate use or non-use of a 15-minute slot of time for Medicaid activities. The advantage of this choice of a dichotomous variable is that the variance follows directly from the choice, and some kind of variance estimate is needed to proceed with a sample size calculation. The choice of a dichotomy is a conservative assumption, with a conservative influence on sample size estimates, as shown below. Although the model proposes to sample five full days, data will be collected on whether or not the staff member engaged in Medicaid related activities for each 15-minute time moment within those five days.

## **The Statistical Sampling Frame for Selecting Staff**

***Eligible staff in small districts:*** The sample selection of all (n=5,199) eligible staff in small districts increases the precision of the statewide estimate and facilitates the equitable distribution of Medicaid funds to individual school districts.

- Smaller school districts vary considerably in the amount of time spent on Medicaid activities: between 1% and 10% or more. Random sampling of eligible staff in smaller districts would generate large random sampling errors.
- Smaller school districts have a higher than average time spent on Medicaid activities: about 6.2 % compared to the 3.7% statewide average. Sampling all staff in these smaller school districts contributes to the accuracy of the statewide estimate.

Furthermore, the allocation of Medicaid funds within Washington State to each small school district is easily calculated based on each district’s results.

***Eligible “designated” staff in larger districts:*** The sample selection of all eligible “designated” staff (n= 2,650- 5,298) increases the precision of the estimate for each larger district and also increases the precision of the statewide estimate, as well as facilitating the equitable distribution of Medicaid funds to school districts.

- “Designated” staff, even though a small proportion of all eligible staff, are estimated to spend a larger proportion of their time than teachers and other support staff. They are involved in more time-consuming types of activities: planning, outreach, helping to fill out eligibility forms, informing parents, connecting with medical service providers, etc. Measuring the time spent by all designated staff decreases the random sampling error of estimates for these larger districts.

Furthermore, the allocation of Medicaid funds within Washington State to each of the larger districts is facilitated by this extra precision. Each larger district will receive funds proportional to the time spent by these designated staff (see example of allocation procedures: Attachment 8).

***Eligible “teachers and other support staff” in larger districts:*** Random samples (stratified by school) are proposed, since this category represents a very large number of staff. The sample universe is estimated to be as large as N=58,550 staff. The recommended sample size for the first quarter is 25% of the sample universe: n=14,640 teachers and other support staff. This is a reasonably conservative estimate of the sample size necessary to reach statistical validity.

Though not fully applicable (due to non-independence of many 15-minutes reported by the same staff, potential non-independence of staff within the same school, potential non-independence of schools within the same district, and potential variability of the parameter P) a formula useful for calculating the necessary sample size for teachers is the following:

$$n = z^2 N P(1-P) / [(N-1)e^2P^2 + z^2 P(1-P)]$$

with

N = 91,338,000 (RMS moments: 58,550 teachers times 1,560 time periods)

z = 1.96 (Two tailed, 95% confidence)

e = .05 (5% error in estimates)

p = .0209 The proportion of 'RMS hits' expected for the sample universe of teachers in larger districts.

Rather than using this equation to justify the sample size, it is used to establish a large sample size under very conservative assumptions. The conservatism (and safety) of the sample size is shown below. In fact, the precision of the statewide estimate is more likely to be 2-3% than 5%.

(The 2001-2002 Washington statewide estimate is .0295 for these larger districts, after including the ten percent reduction in order to account for stricter definitions activities that can be considered Medicaid related ones in the future. This estimate is further reduced by the estimated time of designated staff of about .0086, which gives us the final .0209 figure.)

This formula produces a sample size of 71,932 RMS moments. If five 15-minute time samples are chosen, then the teacher sample size needed is 14,386 ( $5 * 14,386 = 71,932$ ). The figure of 14,386 teachers is very close to the one-quarter fraction of teachers, 14,640, and the operationally simpler one-quarter figure of 14,640 teachers will be used.

#### *Methodology:*

1. Time: randomly select five days (instead of five 15-minute periods) per person for each school in each school district.
2. Persons: randomly select one fourth of the teachers in each larger school district ( $58,550 / 4 = 14,640$ ).

*Statistical note on adequacy of proposed sample sizes of person-days:* The following discussion of the statistical adequacy of sample sizes was recommended by the independent reviewer, Nayak Polissar Ph.D. He suggested statistical analysis that could show directly the conservatism (safety) of the choice of a sample of one-quarter of the teachers.

First, the one-quarter figure arises from the simple random sample format used just above, with the sample size derived from the equation noted. In this Washington school setting, the precision of the estimate from an appropriate *stratified* random sample, such as that which will be used in the proposed time study design, will be superior to that of a simple random sample.

Second, the proposed sample size of one-quarter of the teachers ( $n = 14,640$ ) in large districts, will provide better than 5% precision for the statewide estimate of percentage Medicaid time, even if the

format were to be a simple random sample. The excellent precision can be shown as follows. Just for this demonstration, the assumption is made that a simple random sample will be drawn of teachers in the large school districts. For simplicity, we ignore the finite population correction (This correction leads to greater precision, and by ignoring it in this demonstration, we are understating the precision of the simple random sample.) The precision of the estimated percentage Medicaid time is defined as the half-width of the 95% confidence interval for the estimate, which, for a simple random sample, is

$$\text{Precision} = z * S / \sqrt{n}.$$

As defined earlier,  $z = 1.96$ , corresponding to 95% confidence, and  $n =$  sample size. The quantity  $S$  is the sample standard deviation of the percent Medicaid time, calculated across the sample of  $n$  teachers.

If the mean value of percent Medicaid time is  $M$ , then the precision can be presented as a percentage of the mean by the equation

$$\text{Precision (\%)} = 100\% * z * (S/M) / \sqrt{n}.$$

The quantity  $(S/M)$  in this equation is the coefficient of variation (CV), and in most realistic settings it is under 1.0. (A CV of 1.0 would be considered highly variable data.) The precision (%) equation shows that with the sample size of 14,640, the CV could be as large as 3.1 and still yield a precision (error) of 5%.

A CV of 3.1 is extraordinarily large and is unlikely to be encountered in practice. For example, economic data, of which percent effort is kindred, may be quite skewed and have a large CV. The lognormal distribution is sometimes used to model economic and environmental data, and a CV of 3.1 for the lognormal distribution would imply that the 95-th percentile of the distribution would be about 18 times as large as the mean—extraordinary skewness. If the observed CV is 1.0 or 2.0 in the Washington data, then the precision of the mean percent Medicaid effort would be 1.6% or 3.2%, respectively—under the 5% precision goal.

In summary, while the sample size for the first quarter is based on convenient and usually conservative assumptions, the expected precision is likely to be substantially better than 5%, due to

- a) the use of stratification in sampling,
- b) the multi-day sample per teacher,
- c) the inclusion of 100% samples of designated staff in all districts, and
- d) inclusion of 100% of teachers in smaller districts.

**Proviso:** Recommendations to modify the sampling design will be made on the basis of standard statistical formulas, using estimates of time variances from first quarter results. Possible modifications might include:

1. decreasing the number of days or staff sampled if the day-to-day variance or person-to person variance is small; or
2. decreasing the number of days or staff sampled if the distribution of time is found to be independent of staff/school district characteristics; or

3. increasing the number of days or staff sampled if the day-to-day variance or person-to person variance is large;
4. modifying the method of selection of days in a quarter giving different sets of staff members different sets of randomly sampled days. For example it may be practical in larger school districts to sample different sets of days by type of school (elementary, middle, high school, alternative schools) or by prevalence of Medicaid students (high, medium, low prevalence).

### **Sampling Decisions in this Modified Design Produce Conservative Statewide Estimates**

- Stratification reduces error: days stratified by day of the week and beginning/middle/end of quarter, and persons stratified by school and function in larger school districts.
- Randomly selecting five days instead of five 15-minute periods provides data for 130 moments of time and allows calculations of time variances for future design modifications.
- A two-tailed test of significance at 95% confidence means that the chances of overpayment are actually only 2.5%.
- Decreasing the estimate of the proportion of time spent by teachers and others in large school districts by 10 percent increases the sample size required.
- Among small school districts, and designated staff in large school districts, the only error is associated with time since all staff are sampled.

In summary, the modified stratified design collects information from 23,814 school staff sampled on five random days, involving 130 fifteen-minute periods for a total of 3,095,820 person-time moments.

### **Major RMS Features are Preserved to Provide Unbiased Estimates**

The proposed time study follows major RMS requirements

#### ***Random sampling***

- Random samples of similarly defined eligible staff and time are drawn. (See Attachment 1 on definition of eligible staff and school days)
- Only a minimum notice is given, prior to the study, of who the sampled persons are and when the sampled days will occur.
- Sample universes are defined for each school district, and samples are drawn at random, by an entity 'external' to the school districts: the Medical Assistance Administration staff in the Department of Social and Health Services (DSHS).
- Random samples of persons and days are taken prior to the beginning of each quarter for each participating school in each school district.

***Clear definitions of activities to be measured***

- All staff in the sample universe are equally trained, with the same guidelines and definitions. (See Attachment 5, on training)
- The data collection forms are uniform. (See attachment 1, on data collection form)

***Unbiased data collection procedures***

- Sampled time results are collected soon after they are recorded – day logs are turned in within specified deadlines.
- It is assumed that unreturned forms reflect ‘0 time spent on Medicaid activities’, a procedure that produces bias in a conservative direction (underestimating, not overestimating, average time spent).
- No substitutions are allowed except for staff that may no longer work for the district (left employment, died etc.), or days no longer considered school days (snow days, teacher strike days, etc.). Substitution staff and days are chosen at random by the external entity, DSHS.
- Time study procedures are documented. (See Attachment 4, on documentation)
- Random data quality checks are conducted by DSHS. (See Attachment 6, on monitoring)

***Clear definition of who does what***

- DSHS and School Districts have clear responsibilities and roles (see Attachment 7, on DSHS contracts with school districts)





## **Detailed Features of the Proposed Design**

### **The Sampling Universes of Eligible Persons and Time are Clearly Defined and Verifiable**

The proposed study defines the set of eligible school staff to be included in the sample universe. It excludes about 23% of all school staff through a set of clearly defined program and duty codes maintained by the Office of the Superintendent of Public Instruction (see OSPI codes and statewide and school district staff counts, in Attachment 2: Definition of Sample Universe). These codes exclude staff fully funded by other federal sources, staff whose expenses are covered by the ‘indirect cost’ allocation formula, and staff employed in specific food, transportation and building maintenance programs.

The study includes in the sample universe of time all the officially defined “school days” in the particular quarter. They average 60 days per quarter in the legally prescribed 180-day school year, even though they may differ district by district. Field trip days and parent conference days are also included. The days are officially posted on the OSPI web site before the beginning of each year (see Attachment 2).

### **Calculation of Overall Statewide Estimates Involve Weighted Averages Due to Different Sampling Fractions and Sizes of Strata**

Since in some strata the design calls for sampling all staff, in others it calls for sampling only a quarter of the staff, observations will have to be appropriately weighted in order to provide an accurate statewide measure of time spent on Medicaid related activities.

This can be achieved in a two step process:

1. Weighting the appropriate results for each of the larger districts (combining full samples of designated staff with quarter samples of teachers in each district, using weights proportional to FTEs)
2. Weighting the appropriate summary results from the larger districts with those from the smaller districts, , using weights proportional to FTEs.

### **Allocation of Funds to Individual Large and Very Large School Districts is Based on Two Time Estimates: a district specific one for “Designated” Staff and a pooled one for “Teachers and Other Staff.”**

These calculations do not affect the precision of statewide estimates of time spent on Medicaid related activities. It provides an efficient, and hopefully fair way, of allocating funds internally to the state of Washington, among larger districts.

Ratios of “teacher time” to “designated time” can be calculated based on pooled averages (pooling across multiple districts). Pooled estimates are sufficiently precise for two ratios to be reliably accurate: one for the 13 largest districts, the other for the 93 other large districts (see Attachment 9 which provides a bar chart depicting the frequency distribution of school districts by size of staff).

Currently there is no information on the variability of such ratios across larger school districts. This will be calculated after the first quarter results are obtained.

The two time estimates, the ‘designated’ staff one and the ‘teacher’ one, are expected to be highly correlated if either, or both, of two conditions apply:

1. The higher the time spent by teachers referring and monitoring students, the higher the time spent by designated staff connecting students to Medicaid.
2. The higher the time spent in program development, planning and training and in outreach activities by designated staff, the higher the time spent by teachers in referring and monitoring students.

The magnitude of this correlation can be calculated after the first quarter results become available. If the correlation is relatively high, the residual standard error is relatively low, and the ratio is fairly constant across districts, then the calculated ratio from pooled data can be used to allocate the pooled time for “teacher and other staff” to each larger school district.

This is accomplished by multiplying the calculated ratio (pooled teacher time/pooled designated time) times the measured “designated” staff time unique to each larger school district (See the illustrative example of pooled based calculations in Attachment 8). These calculations are done separately for 93 large and 13 very large districts.

### **Sample Sizes in the Modified Design Compared to Those Required in a Simple Random Sample RMS Design**

In order to calculate the necessary sample size to meet the requirements, it is important to identify whether the variable “time spent on Medicaid related activities” can be considered a dichotomous (yes/no) variable or a continuous one (amount of time or % of total time available per staff). This influences what formulas to use. (The discussion, earlier, about precision of estimates from the sample for large school districts includes comments and qualifications that apply equally to this discussion.)

If we measured the presence of Medicaid related activities for individual 15- minutes time periods, the variable would be dichotomous: either staff did or did not spend those fifteen minutes on Medicaid activities. The data collection procedure would entail asking staff about particular, randomly chosen 15- minute periods. Under these conditions the following formula could be used to calculate the necessary number of RMS moments to be sampled:

$$n = z^2 N P(1-P) / [(N-1)e^2P^2 + z^2P(1-P)]$$

with

N = 103,573,000 (RMS moments)

z = 1.96 (Two tailed, 95% confidence)

e = .05 (5% error in estimates)

p = .03366 (The proportion of ‘RMS hits’ expected for the sample universe)

This is the 2001-2002 Washington statewide estimate of .0374 reduced by ten percent, in order to account for stricter definitions of activities that can be considered Medicaid related ones in the future.

This formula produces a sample size of 44,608 RMS moments. If five 15-minute time samples were chosen, and a different sample of staff completed each 15-minute time sample, then the staff sample size needed would be 8,921 ( $5 * 8,921 = 44,608$ ).

This simple random sample RMS design involves a much smaller number of staff than the modified design proposed here: 23,814 staff (5,199 small school district staff, plus an expected average of 3,975 designated staff, plus 14,640 teachers).

*Statistical notes on adequacy of the sample size for statewide estimates.* The situation here is similar to that for the discussion of the sample for large districts (earlier in this document). The sample of 14,640 teachers would be expected to give better than 5% precision (error) from any set of districts, as noted in the earlier discussion. In the proposed design, a substantially larger total statewide sample size is proposed than that required by a simple random sample RMS design, and precision will be, correspondingly, even better than noted earlier. Precision is, again, expected to be considerably better than 5%.

Given the differences between a simple random sample design and the modified design proposed here, it is useful to recall which modifications were made and why.

### **Summary of Reasons for Modifying the Simple Random Sample Design**

*Practical reasons: clustering time* – In practice, it is difficult to measure time spent in a particular fifteen minute moment (interrupting school staff engaged largely in teaching and other educational activities). Furthermore, staff would get annoyed and become non responsive since they would have to respond ‘no’ most of the time. Aggregating moments into clusters of days is a much better alternative. A daily time log of activities is much less disruptive and can collect more information more efficiently.

*Characteristics of schools and staff: stratifying by district and by staff* – Staff in smaller school districts, often in more rural and poor areas of the state, spend more time in Medicaid related activities. Staff in these smaller schools ‘wear many different hats’: they do not specialize in what they do. Many different types of staff engage in Medicaid related activities. Past data suggest stratification of schools in two strata: 99 smaller school districts, 106 larger ones.

In larger school districts, on the other hand, exploratory research suggests that few designated staff engage in the more time consuming Medicaid activities (outreach, compilation of forms, contacting parents and providers etc.) while teachers mainly identify, refer and follow up students. This suggests stratifying staff in larger districts into two strata: designated staff (comprising 4-8 percent of total staff) and the rest, mainly teachers.

*Allocation of funds to individual districts* – The plan in the Medicaid Administrative Match program is to continue contracting with each school district. The districts want to be reimbursed based on the magnitude of their specific efforts. No demographic or school characteristics have been found to predict individual district time accurately enough. Only 30 percent of the variance can be explained by a variety of such factors. Therefore, the sampling design needs to generate samples large enough to estimate the time spent by each individual school district.



## References

- Bolstein, R., G. Ballentine, J. Vaughan. 1986. "Implementing Random Moment Sampling in the Virginia Department of Social Services: Design and Analysis." *Proceedings of the Twenty-Sixth Annual Workshop of the National Association for Welfare Research and Statistics*.
- Levy, P.S., and S. Lemeshow. 1991. *Sampling of Populations: Methods and Applications*. N.Y.: John Wiley and Sons, Inc.
- Niebal, B.W. 1972. *Motion and Time Study*. Richard D. Irwin, Inc.





