APPENDIX D

SUMMARY OF THE DESIGN OF THE RESEARCH

Summary of the Design of the Research

I. Trends and Projections

The analysis of trends and projections applied different types of statistical analysis to aggregate data (weekly or yearly) provided by the various justice agencies, in most cases covering at least a ten year period ending in 2005. The first step in the trends/projections analysis was to model statistically the aggregate data over time, through a variety of time series techniques (including curvefit and exponential smoothing) to determine the models with the best fit for the data. In the next step, projections were carried forward, usually for a short period given the limitations of the data (for projections to be reliable, the projected period should not exceed a quarter of the pre-projection period (modeled data). The trends models selected were in each case among the most conservative of model choices. The resulting projections, therefore, were not the most extreme possible under some statistical models (not selected) and, in fact, may have produced under-estimates in some cases. In other words, other model choices would have resulted in more dramatic increases in the trends we projected. Our conservative stance in projection is because all projections are based on past data and carry with them the unrealistic assumption that the future will be mostly like the past projected forward. (In our statistical experience, the future is often not like the past in all sorts of ways once it becomes known in the present.) However, trends and projection analyses are mainly helpful as the basis for planning discussions that consider whether or not assumptions like those shaping the past will also shape the future.

It is possible that various sources of the aggregate data analyzed in this report carry with them certain limitations (like missing data) and inconsistencies. For example, improvements in reporting or verifying data may mean that totals over time are measured somewhat differently—and that changes in recording procedures themselves produce changes in the levels and numbers of what is being measured. Although they are the best available for planning purposes, these possible limitations represent another reason for making use of these data with caution.

II. The Prison ("On-a-given-day") Snapshot Study

In employing the single-day approach, we recognized that the overall size and characteristics of the population of the Prisons may change from day to day and fluctuate over time (during the week, month and year). (For an illustration of this, see Figures 1 and 2 in Appendix A for Chapter Two, which chart the upward trend in the annual average daily population of the Philadelphia Prisons from 1960 through 2005.) This dynamic property of the inmate population is illustrated, for example, by the fact that the date selected for study last November preceded the transfer back to Philadelphia of more than 300 inmates from Delaware County. Seasonal patterns are also well-recognized by local officials (and they were apparent in the trends analyses) (e.g., after August vacations, during end-of-the-year holidays, etc.) and affect the possible make-up of the population on a given day.

This limitation aside, however, this single- or typical-day approach can nevertheless provide an informative, cross-sectional ("snap-shot") look at the composition of the Philadelphia Prisons of value in planning for both institutional and community-based correctional capacity.

Minor variation in population composition aside, an in-depth examination of one day can reasonably well represent how the population might look on many days and point to features of the processes that contribute to it.

The findings presented in Chapter Three are based on an in-depth study using both overall population data from the Prisons computer system (N=8,541) and a sample (n = 700) carefully drawn to represent the entire inmate population on November 21, 2005 (with the exception of "out of custody" inmates, N=8,415). Where data were available for all inmates, we employed full population data (this was rare). For most of the analysis, considerable in-depth data collection and source cross-checking was involved; this was done for the sample of inmates. Data were collected relating to legal, case-related, demographic, custody-related and other attributes associated with each of the sample inmates. Except where noted, the percentages reported are sample percentages and, as such, should be interpreted as estimates of the total population of inmates on that day and as having a small margin of error around the actual population values.

The sampling approach involved disproportionately stratifying by institution: random samples of 100 were drawn from seven institutional categories to guarantee inclusion of all inmate categories and to minimize standard error. These were weighted based on the inverse of their sampling fraction to produce estimates of the full inmate population, plus or minus a margin of error. The following table provides illustrations of the standard error and confidence intervals associated with different size estimates of inmate population attributes. Note that for all estimates, whatever the split of the population on a given attribute (e.g., 10/90, 20/80, 30/70, 40/60, 50/50), the same standard error applies to each side of the split. For example, the estimates for gender are split 9.9 percent female and 90.1 percent male with a standard error of 0.3 percent. This margin of error would apply to each of the gender estimates to determine their respective true value ranges.

Examples of standard error ranges surrounding sample estimates reported in Chapter 3, at the 95 percent confidence level (Confidence Interval = Estimate +/-2Std.Errors)

Variable	Estimate		Std. Error	Population Value Range	
	%	N		%	N
Gender (female)	9.9	834	0.3	9.2-10.6	777-889
Prior convictions for weapon offenses	18.8	1580	1.7	15.4-22.2	1292-1872
Prior convictions for property offenses	28.4	2394	1.9	24.7-32.1	2081-2699
Inmates reporting using drugs/methadone	41.0	3452	2.0	37.1-44.9	3119-3782
Prior misdemeanor convictions	51.1	4299	2.1	47.0-55.2	3957-4643

N Population = 8415

III. Processing and Pretrial Release

Because collecting in-depth individual and case outcome data on all (12,333) defendants entering the courts during the March-May 2005 period would be logistically impossible within resource constraints, this study also relied on sampling. The sample drawn to represent all entering defendants during this period employed a disproportionate stratified random sampling design, sampling equal numbers of cases (n=200) from each of the four "zones" of the pretrial release guidelines classification (200 each from ROR, Special Conditions I, Special Conditions II, and cash bail categories). Note that the sample employed the guidelines classification as stratification criterion – as a good way to include all types of entering cases; it did not sample on the basis of guidelines decisions made.

The stratification ensured a) that estimates for a full cross-section of all entering criminal defendants would be achieved, and b) that all types of cases would be included in the analyses, not only the most numerous categories, because an important focus was on pretrial release and the guidelines. When weighted for disproportionate selection, the total sample, n=800 cases, produces estimates of attributes and outcomes for the 12,333 defendants entering processing during the spring of 2005. The defendants and their cases were followed for one year from the date of preliminary arraignment, considered as the starting point of the justice process. All samples produce population estimates surrounded by a margin of error, the size of which depends on the size of the attribute reported and the sample/stratum size. For example, the sample suggests that about 20.6 percent of the estimated 12,333 entering defendants during the study period (or an estimated 2,542) had serious personal charges. The margin of possible sampling error at the 95 confidence level can be calculated as follows: With two standard errors equal to 2.3 percent, we can state that the true population value (based on the 12,333 defendants) falls between 18.3 percent (2,254 defendants) and 22.9 percent (2,872 defendants).

The following table provides additional illustrations of the standard errors and confidence intervals associated with different size estimates of the attributes of the population of 12,333 defendants entering the courts during March-May 2005. Note again that for all estimates, whatever the split of the population on a given attribute (e.g., 10/90, 20/80, 30/70, 40/60, 50/50), the same standard error applies to each side of the split.

Examples of standard error ranges surrounding sample estimates reported in Chapter 4, at the 95 percent confidence level (Confidence Interval = Estimate +/-2Std.Errors)

Variable	Estimate		Std. Error	Population Value Range	
	%	N		%	N
VUFA charge as most serious charge	2.2	272	0.5	1.1-3.3	137-406
Any VUFA charge (whether or not leading)	8.0	986	1.0	5.9-10.0	736-1237
Deviations from guidelines because of prior criminal history	11.1	1364	1.1	9.0-13.2	1106-1632
Any serious personal charges (whether or not leading)	20.6	2542	1.2	18.3-22.9	2254-2827
Prior arrest (one or more)	68.4	8432	1.7	65.1-71.7	8029-8842
Prior felony arrest (one or more)	59.0	7280	1.7	55.6-62.4	6856-7697
Leading charge type – Felony	55.5	6845	1.2	53.1-57.9	6545-7145
Deviations from guidelines (Yes)	48.4	5972	1.8	44.9-51.9	5542-6396

N Population = 12333