The Legacy of Data in Shaping Homelessness Policy[1]

Walter Leginski, Ph.D.
Office of the Assistant Secretary for Planning and Evaluation
Department of Health and Human Services

America was reintroduced to homelessness during the 1980's.

The country had not seen an epidemic of homelessness since the Great Depression and in the 1980's, new generations of citizens were exposed to homelessness on the streets of their communities, on the covers of national magazines, and in popular entertainment such as made-for-TV movies. Initially, we did not understand its causes, why homelessness was suddenly so visible, who was becoming homeless, nor how to channel the compassion, guilt, and anxiety that homelessness evoked. As we look back, nearly every feature of our current legislative and service response, our understanding of the epidemiology of homelessness, and our knowledge about effective solutions are a legacy of the responses that were begun in the 1980s.

This retrospective is apt. The panel representatives today are individuals who have been involved in a variety of efforts focused on homelessness data collection. Particularly important to the panel has been the challenge of proposing what might be the nature and content of a national data collection during this decade. We need to consider what value prior efforts have yielded, what new technologies and policies have emerged, and the degree to which we build on or depart from our last major national effort in this area, the National Survey of Homeless Assistance Providers and Clients (NSHAPC: Interagency Council on the Homeless, 1999).

From the earliest stages of emerging awareness about the new wave of homelessness, there were organized efforts to understand the problem of homelessness, either by putting a number to it, by identifying the characteristics of those who were homeless, or by conducting research on effective service delivery. Many of these early efforts were local in nature—usually at a city level and often driven by an advocacy agenda (Culhane, 2001). By the mid to late 1980's, survey efforts were increasingly systematic and representative. Examples include:

- in 1985, the U.S. Conference of Mayors began its annual report on hunger and homelessness in more than 20 major cities (U.S.Conference of Mayors, 2002);
- the 1986 the Ohio Department of Mental Health released findings from a State-wide survey of homelessness (Roth & Bean, 1986);
- the Department of Housing and Urban Development commissioned the first national survey of homeless shelters (Dietz, Light, & Marker, 1984); and

---

[1] The views and opinions expressed are those of the author and do not reflect official policies of the Department of Health and Human Services.

- the Urban Institute conducted the first nationally representative survey of homelessness in 1987 (Burt & Cohen, 1989).

As we consider these early efforts, one fact that is striking is the speed with which findings from these undertakings were transformed into legislation, policy, resources and programs. In this paper, I suggest that this established a pattern that has persisted for more than 20 years within the community of agencies, organizations and researchers who focus on homelessness—a pattern that generally values the rationality of using empirical findings to inform legislation, policies, targeting of resources and the focus of programs. As we think about future data collection efforts, we have an opportunity to continue or change this pattern. That opportunity will be also be addressed in the paper.

A few examples will illustrate the tenet that traditionally there has been an explicit transition from data to policy and action in the area of homelessness:

- By the late 1980's, governments and providers had conducted sufficient research to describe the composition of the homeless population in ways that have contemporary vestiges and accuracy. Specifically, by 1989 (Tessler & Dennis, 1989), the high prevalence of behavioral health problems (substance abuse and mental illness) and primary health conditions was well documented. It is not uncommon to still see references that among the homeless population, one-third have a psychiatric problem and half have a substance abuse problem—figures that have their base in the documentation done in the mid to late 1980s.
- By the late 1980's Congress had established a targeted block grant for States to serve homeless people with mental illnesses. In 1991, this became a formula grant program to States known as PATH (Projects for Assistance from Homelessness). PATH continues to be funded and managed by the Department of Health and Human Services (HHS) and in FY 2003, it received an appropriation of $43 million.
- The prevalence of physical health problems among homeless persons and the problems they had accessing even public clinics led foundations to respond. Most notably, in 1985 the Robert Wood Johnson Foundation and the Pew Memorial Trust sponsored a demonstration of health care for the homeless. Within 2 years, this evolved into a targeted program in HHS—Health Care for the Homeless—that currently supports more than 150 grants with an appropriation of approximately $120 million in FY 2003.
- By the late 1980's, outreach workers—those who went on to the streets to find and engage homeless people—were no longer an innovation. Demonstrations in the mid 1980s had posed the question: Would outreach engage homeless people and help move them from the streets? By 1990, outreach workers were an identified personnel specialty in the field and continue to be key players in homelessness services (Interagency Council on the Homeless, 1991).
- In the late 1980s to early 1990s, innovations in the administration of homeless service programs in many communities led to centralization of

intake and coordination of the activities across providers. New databases allowed better counts of service users and a description of service use over time. User groups were formed to explore common analytic approaches and offer peer-to-peer technical assistance. The National Human Resources Data Consortium (www.nhsdc.org) is the most active of these.

- In 1998, Kuhn & Culhane used homelessness administrative databases to confirm distinct subgroups of homeless service users—a taxonomy that had first been suggested by the Institute of Medicine (IOM) in 1988 (IOM, 1988). Looking at service use over time and across several cities, they confirmed the IOM suggestion that the population can be categorized into 3 unique groups—those temporarily, episodically, or chronically homeless. Within 2 years the National Alliance to End Homelessness had implemented a campaign to eliminate chronic homelessness in a decade and in 2002 this target was adopted by the Bush administration in the budget for the Department of Housing and Urban Development.

As we think about the contributions that data have made to the accumulation of homeless legislation, policy, resources or programs, we have an opportunity to ask:
--what characteristics of the data contributed to their impact? and
--which of these characteristics should we retain and replicate in the formulation of future data collection on homelessness?

Nine data collection characteristics are offered for consideration. For each of these characteristics, we can assign a 'score' that tells us how significantly it contributed to legislation, policy, resources or programs (LPRP). If the characteristic contributes in a major way, it receives a score of 10. If its contribution in one of these areas appears negligible or is offset by other considerations, it receives a low score. Three caveats are important. First, the scoring is a heuristic device to stimulate our thinking about future survey work. Second, it is subjective in that it draws on the experiences of the author from his perspective in Federal government. Third, the scores are based on impacts on LPRP. Judged by different impact criteria, say from the perspective of a homeless person or a homeless service provider, the characteristic might be assigned a different score.

1) Importance of Federal Sponsorship
Score: 3
Rationale: The chief value of Federal sponsorship is financial—Federal agencies are most likely to be able to fund extensive surveys. But there is no consistency of sponsorship in our history to suggest this is the best approach. Local, State, foundation, association, university, and Federally sponsored data collection has had detectable influence on LPRP.

2) An Estimate Of The Size Of The Population
Score: 10

Rationale: This will always be a controversial and elusive endeavor, but it remains fundamentally important. In the 1980s it was critical because it helped us understand that this was not an isolated, small phenomenon, but rather that larger social forces were in play. The size of the problem triggered action. In designing the 1996 survey, this dimension was considered, but other policy factors were given higher value and a size estimate was not included as a design feature in that Federal undertaking.

However, in the 2000s, a size estimate is again important because of the new focus on ending chronic homelessness by 2012 and by the emerging trend of cities to issue 10 year plans to end homelessness as reflected in reports from the U.S. Interagency Council on Homelessness (www.ich.gov). Documentation of the success of these efforts over time will be critical.

3) Representativeness of the Data

Score: 9

Rationale: The first cut at representativeness would be for the data to be nationally representative. This is generally valued and desired, but it is not a core requirement. An estimate or finding based on a large N is also valued. Having a large number of observations suggests a stability in the finding and its potential applicability elsewhere. For example, a study of the cost offsets of placing mentally ill homeless people in permanent housing in New York City that involved more than 800 subjects (Culhane, Metraux, & Hadley, 2002) is widely used by other cities and States to describe what they can expect (see State action plans at www.hrsa.gov/homeless). It is also assumed to apply to other chronically homeless subgroups and to reflect cost savings that can be expected, even when their service use profile is likely to be different.

4) Composition of the Population

Score: 10

Rationale: Being able to report on the make-up of the homeless population is the most consistent reason data collection on homelessness has so readily been translated into LPRP. Particularly important has been the documentation of major health and behavioral health conditions, the representation of families with children among the homeless population during the last decade, the high representation of veterans/combat veterans, and the ability to describe subgroups based on their use of services or length of time homeless.

As noted above, composition data has often led to the development and funding of targeted responses such as PATH, homeless veterans programs, and a focus on ending long term and recurrent homelessness. Composition of the population will continue to be a critical information need that must be filled by data collection on homelessness and may be the driving rationale.

5) Timeliness of the Data

Score: 6

Rationale: So far, homelessness at the level of legislation and policy has not shown a need for real-time data such as is required during a disease outbreak, like SARS or the West Nile Virus. Homeless assistance programs do prefer data that is close to real-time. They need it to make resource allocation decisions such as where an emergency shelter bed, transitional housing, or food program may be available But surveys are not the best way to provide these data to programs and HUD has been supporting the development of on-line, real-time information systems to support this need (HUD, 2002).

There does appear to be an advantage in producing data that refer to a recent time interval, but the homelessness field accepts as useful data that other fields would view as dated. The data from the last national survey on homelessness took 3 years until release without complaints about its applicability. Analyses of administrative data have looked back 3-5 years from current time periods and the findings have a high degree of acceptance. The U.S. Conference of Mayors produces homelessness data in December of each year (2002). These annual reports are a valuable source for looking at trends, e.g., family homelessness has remained around 40 percent of the annual requests for help, but the speedier cycle of these data has not yielded noteworthy annual changes. Therefore, within some reasonable time period, recent data are valuable, but real-time is not essential.

6) Continuity of Data Content

Score: 9

Rationale: There are many advantages to continuity in the content of the data over time. Continuity is especially important when a data collection effort is repeated, even when the time intervals are widely spaced. Our national decennial census is an example, where the ability to track changes over time in measures that remain constant yields great value. Another version of continuity is evident when similar content is addressed in different data collection efforts. In the homelessness field, a focus on health conditions, family status, and homeless histories in many different data collection efforts tends to reveal repeated patterns that validate findings and allow us to track changes. In the design of the 1996 NSHAPC, considerable attention was given to continuity with the Urban Institute study from the late 1980s. However, circumstances, needs, utility, understanding, and technology change and require that opportunity for revisions to content be encouraged. The more widely spaced the data collection intervals, the more important it will be to allow for content changes that capture new circumstances.

7) Point in Time or Longitudinal

Score: 5

Another time-focused consideration is the length of time covered by the data collection. Most data collection focuses on a specific point in time, asking providers or clients who have been explicitly selected to report on a set number of characteristics. Most often these approaches draw on recollection (e.g., Link et al., 1995) or the embedding of instruments within the survey that allow for the calculation of an index, particularly health functioning, as was done in the

NSHAPC. These approaches have been particularly important in fulfilling the composition feature noted above.

A more recent approach allows for longitudinal analysis by drawing mostly on the use of administrative records. In this case, longitudinal refers to the use of information systems to construct linked data records on providers and clients, not the tracking of a cohort of individuals through some period of maturation or experience. Constructing a longitudinal set of data from service use allows us to apply analytic techniques based on time series. Longitudinal approaches also contribute to an understanding of composition, with the added value of being able to compute survival analyses and associate specific activities, such as amount or types of service with providers and clients.

Both approaches are useful and inform us about different aspects of composition and a data collection effort on homelessness must generally choose one of these approaches. However, the choice is based on considerations other than influence on LPRP, such as cost, feasibility or the questions to be examined.

8) Quality of the Endeavor
Score: 10
Rationale: This refers to the quality of the methodology and the assurances we can derive that the findings have credibility and generalizability. Four dimensions are key. They are sufficiently interdependent that it is appropriate to consider them together rather than as separate characteristics.
First, *selection criteria*. This can be based on a formal sampling frame to select those covered by the data collection or a complete data universe, as might be the case in an administrative data analysis. Convenience samples come with many limitations to the usefulness of the data.
Second, *quality of the execution*. This covers such issues as the training and professionalism of the data collection staff, informed consent, quality control procedures, and the efforts made to ensure that unbiased, uniform, and objective data were obtained.
Third, *psychometric properties* of the measures used. Some attention must be given to data collection that provides reliable and valid information. Instrumentation that has been used previously or which has had its psychometric properties established elsewhere are the typical ways this is done.
Fourth, *analysis of the data*. The use of appropriate statistical techniques to produce numbers and percents must consider sampling, weighting, mode effects and other aspects of the data collected. Tests or statements of significance must be considered against the data and the assumptions it can and cannot support.

To some extent, the more formal the sponsorship of the data collection, the more likely quality factors will be considered. In the conduct of Federally sponsored data collection, each dimension receives heavy attention. This scrutiny on quality ensures that the resulting information will be trusted as authoritative and representative. Quality has resulted in considerable impact on LPRP.

Where these factors are most readily applied is in designing and conducting surveys and special studies. Where they can be difficult to apply or fix is in the analysis of administrative data. This is of concern when sources such as homeless management information systems (HMIS) become the exclusive source of national data on homelessness. Analyses of such administrative data offer many advantages, but as multiple administrative files are accumulated across multiple settings, many researchers find problems.

--They may find they are combining both sample and universe data.

--The data may be available only from a nonrepresentative set of providers. For example, if the data are from providers who volunteer to contribute data, have a unique technical capacity to produce a file, or are paid to provide the data, it can be challenging to know what population of providers or clients is represented and to what extent findings are generalizeable.

--If the efforts are in one geographic area, the same person may be counted more than once as you accumulate records across multiple providers. Since homeless individuals can be heavy users of services—e.g., moving from a shelter to a food program to a health service program to a drop in center to a job assistance program all in one day—such a pattern can substantially affect the data.

Even when administrative databases have been used for authoritative reporting, such as to legislative or funding bodies, problems in the data quality are not unknown. As such administrative data bases are reexamined or combined with other data sources, incomplete and missing data, keying errors, inconsistencies in the categories used for particular items, and incompatible time frames will be encountered (Leginski et al. 1990).[2] Fixing some these may require dropping legitimate observations or reconfigurations of the data that can be challenged on the basis of reliability and validity. Unless there is regulation or industry standardization affecting these systems, the quality problems associated with administrative data can be challenging. Consequently, an over-reliance on such data to impact LPRP should be avoided.

9) Cost of the Undertaking

Score: 5

Rationale: There does not appear to be an association between the financial investment in the data collection and its impact on LPRP. The 1996 NSHAPC was a mutli-million dollar undertaking supported by 12 Federal agencies. It effects on LPRP were no greater than foundation sponsored analyses of cost offsets in New York City that cost a few hundred thousand dollars.

---

[2] One supplier of data was prevented by statute from eliminating records from a database if the patient had eloped from the institution. No adjustment for this circumstance was made when using the data internally. When the database was used for multi-site investigation, it revealed that clients with dates of birth in the 19th century and/or ages in excess of 100 years had routinely been included in reports.

Table 1 summarizes these 9 data collection characteristics and the scores for their impact on LPRP.

Table 1. Data Collection Characteristics of Homelessness Surveys and Their Contribution to Homelessness Legislation, Policy, Resources or Programs

Data Collection Characteristic	**Contribution Score**
Federal sponsorship	3
Estimate of population size	10
Representativeness/Large N	9
Composition of the population	10
Timeliness of data	6
Continuity of data content	9
Point in time vs. longitudinal	5
Quality of the endeavor	10
Cost	5

If we approach a subsequent data collection on homelessness with the above scores in mind, it would appear that if we wish to undertake an effort which would contribute to the development of LPRP, we should consider a

- high quality undertaking,
- with a large, representative N,
- capable of producing estimates and descriptors of homelessness, and
- which had continuity with prior efforts.

It is not inconsequential that this list of high-impact characteristics sounds a lot like another Federally sponsored national survey. However, as noted earlier, this approach is primarily offered as a heuristic device to stimulate our thinking, not as a mandate. The other panelists will enrich our thinking on the usefulness and methodology of getting homelessness data. But the approach does recognize that there is a legacy from the data collected since the early 1980s and that there have been significant contributions from this legacy that deserve consideration.

There is a concluding observation to offer about using such a legacy approach. Much of what drove the accumulation of data and the development of legislation, resources, and targeted programs stemmed from the view of homelessness as a unique circumstance and of “the homeless” as a unique population. All levels of government attempted to understand and manage the problems of homelessness. Governments and providers needed to analyze this phenomenon and explore options for how to manage and address it in humane, efficient ways.

Our discussion today continues in this vein and it is one of the downsides of approaching a future endeavor by tying it to a legacy. These data have also taught us that there are other ways to look at the phenomenon. For example, the population is not homogeneous and it may be better to approach homelessness as an experience rather than a static population. People move in and out of homelessness and understanding the dynamics of this situation may be more helpful to future legislation, policy or resource allocation than a refreshed description of the homeless population.

In particular, one factor that is relatively constant among those who experience homelessness is that they are current or former clients of programs that assist poor and disadvantaged groups. By looking at the dynamics of who becomes homeless or who becomes homeless repeatedly, a different approach is possible *and* is emerging.

This is based on thinking that we may be able to prevent a number of the homelessness incidents from occurring among the clientele of health and human service providers. This might be done through more emergency intervention, flexible funding, mediation services, respite services, improved transition from custodial care to community based care, etc. If our thinking were to shift from how we understood and managed a population that was homeless to how we prevented homelessness in an at-risk group, this would completely alter our thinking about the need, scope, and usefulness of data collection on homelessness itself.

Instead, our focus might be on measuring the performance of mainstream services by viewing homelessness as a bad or negative outcome of such services. Value would be placed on how well they help their clients maintain housing stability and prevent housing loss. Where housing loss did occur, rather than focus exclusively on describing and analyzing the homeless clientele, data collection and analysis would be focused on trying to parse out the contributions made by the provider, the client, and mediating variables. The intent would be to improve outcomes by eliminating actions that led to homelessness or assisting those clients who were most at risk.

Our panel discussions are necessary because there is not yet evidence that the homeless experience is declining in prevalence. But it is useful to remind ourselves that if we thought about homelessness differently, it could lead us to a completely different focus in what we measured, from whom we got data, and what systems and behaviors we were trying to affect with our legislation, policies, resources, and programs.

REFERENCES

Burt, M.R. and Cohen, B.E. America's Homeless: Numbers, Characteristics, and Programs that Serve Them. Urban Institute: Washington, DC, 1989 (UI Report 89-3).

Culhane, D. Responding to homelessness: Policies and politics. Paper presented at the University of California at Berkeley, Institute of Governmental Studies, Conference on Urban Homelessness and Public Policy Solutions. Berkeley, CA. January 22, 2001. http://www.igs.berkeley.edu/events/homeless/

Culhane, D., Metraux, S., and Hadley, T. (2002) Public service reductions associated with placement of homeless persons with severe mental illness in supportive housing. *Housing Policy Debate*, 13(1): 107-163.

Department of Housing and Urban Development (HUD). Homeless Management Information Systems: Implementation Guide. the Department: Washington, DC, 2002 (www.hud.gov/offices/cpd/homeless/hmis).

Dietz, D., Light, H., and Marker, D. Survey Of Shelters For The Homeless: A Contract Report to the Department of Housing and Urban Development. Rockville, MD: Westat, 1984.

Fischer, P.J. & Breakey, W.R. The epidemiology of alcohol, drug, and mental disorders among homeless persons. *American Psychologist,* 46 (11):1115-1128, 1991.

Institute of Medicine. Homelessness, Health, and Human Needs. National Academy Press: Washington, DC, 1988.

Interagency Council on the Homeless. Reaching Out: A Guide for Service Providers. The Council: Washington, DC, 1991.

Interagency Council on the Homeless. Homelessness: Programs And The People They Serve—Findings From The National Survey Of Homeless Assistance Providers And Clients Technical Report. The Council: Washington, DC, 1999.

Kuhn, R. and Culhane, D. (1998) Applying cluster analysis to test a typology of homelessness by pattern of shelter utilization: Results from the analysis of administrative data. *American Journal of Community Psychology*, 26(2): 207-232.

Leginski, W.A., R.W. Manderscheid, and P.R. Henderson. Patients Served in State Mental Hospitals: Results from a Longitudinal Data Base; in NIMH, *Mental Health, United States 1990*. R.W. Manderscheid, and M.A. Sonnenschein, eds., DHHS Pub. No. (ADM)90-1708. Washington, D.C.: Supt. of Documents, U.S. Government Printing Office.

Link, B., Phelan, J., Bresnahan, M., Stueve, A., Moore, R., & Susser, E. Lifetime and five-year prevalence of homelessness in the United States: New evidence on an old debate. *American Journal of Orthopsychiatry*, 1995, 65(3): 347-354.

Roth, D., and Bean, J. (1986). New Perspectives on Homelessness: Findings from a Statewide Epidemiological Study. *Hospital and Community Psychiatry*, Vol. 37, No. 7, 712-719.

Tessler, R.C. and Dennis, D.L. A Synthesis of NIMH-Funded Research Concerning Persons Who Are Homeless and Mentally Ill. Bethesda, MD: National Institute of Mental Health, 1989.

U.S. Conference of Mayors. A Status Report on Hunger and Homelessness in America's Cities. The Conference: Washington, DC, 2002