

# ***“Many are called, but few are chosen”:* Appraisal Guidelines for Sampling and Selecting Case Files**

by **TERRY COOK**

This article outlines methods for choosing “representative” or “exceptional” groupings of case files from a larger series.<sup>1</sup> The aim of all such methods is to preserve the significant value of the whole in the part sampled or selected for archival retention. The article begins by introducing the problem of case files and by briefly mentioning the broader theoretical and strategic issues involved in their appraisal. It then considers the characteristics of case files that archivists must analyse before choosing an appraisal option. After reviewing several such options, the article then explores three in considerable detail: formal statistical sampling and exemplary and exceptional selection. Finally, brief sections address the special characteristics of sampling electronic records and the impact of sampling on archival description.

## ***Introduction: The Dilemma of Case Files***

Case files are the most voluminous and routine documents produced by modern bureaucracies. In governments, businesses, universities and similar corporate bodies, they fill records offices and records centres to the brim. If acquired, they threaten to overwhelm archives everywhere with mountains of paper. Their electronic counterparts, while less bulky, also present complex problems. Yet within this avalanche of information are many gems which enrich our understanding of the past. Indeed, such gems can be a sparkling reflection of the citizens’ voice, individually and collectively, and sometimes they are the only such reflection that survives for posterity. Without the patterns and themes uncovered by research in such records, the history of institutions could be told, but not that of people.<sup>2</sup>

Retention of case files by archives may also be essential to protect the legal rights of citizens. Recent examples include the land claims of indigenous peoples, compensation for victims of wartime excesses, or exposure of illegal or unethical intrusions of the state into citizens’ lives (secret brainwashing experiments, contamination of unknowing civilians by nuclear or chemical waste, unacceptable police or security intelligence methods, and so on).<sup>3</sup> Archival retention of case files may also be important for the development and continuity of public policy; in electronic form especially they provide the longitudinal and demographic data necessary to assess the need for change in accepted policies, programmes and attitudes.<sup>4</sup> In certain limited areas (registered

patents for land grants or inventions, etc.), the retention of case files can be essential to the long-term administration of operational programmes. On a broader level, the impact of taxation policies, economic subsidies and research grants may often be assessed through an analysis of such records.<sup>5</sup> Finally, case files are the lifeblood of genealogical research. As people search more and more for their roots, archives will be under pressure to retain more case records to respond to this need.

Yet until recently, archivists have been reluctant to acquire hard-copy case records: "... traditionally, case files have not been retained by government archivists: policy and operational files, with a token sample [*sic*] of case records, have usually been deemed sufficient documentation for any agency." Similarly, the archival acquisition of electronic case records from database management systems has not been undertaken by more than a handful of archives. Even those few archives with electronic records programmes have tended to focus primarily on statistical and survey information. The foregoing research possibilities of case records, therefore, combined with new ways of manipulating the information with the computer, "challenge archivists to define anew their acquisition and selection criteria."<sup>6</sup>

Such redefinition is not easy. In case files, archivists face overwhelming volumes of records — in a modern government context, tens of millions of files are generated annually in hundreds of separate programmes located in thousands of offices. As a result, local variations of practice and procedure enter the records-keeping operations at the case file level, undermining the intended homogeneity of related series of case records<sup>7</sup> and thus rendering more unlikely any statistically valid sampling of the series. Archivists also face the fundamental tension between the archival retention (and resultant public use and cross-linkage) of case files containing personal information and growing concerns about the protection of personal privacy.<sup>8</sup> In addition, archivists face considerable political pressure, through genealogical and other lobbying groups, to retain more case records than they would if archival concerns alone were involved. Recent examples include the celebrated Federal Bureau of Investigation court case in the United States, the redundant land records destruction protests in Ontario, and the attack on archivists in the press and their interrogation before a royal commission over allegations that sensitive immigration case files had been improperly destroyed.<sup>9</sup> Despite such pressure, archivists are also well aware that, in an age of restraint, the retention of unnecessary series of records is not possible. Finally, archivists face this immense problem of appraising case files armed with traditional appraisal theory and often passive acquisition strategies, both of which are quite inadequate to the task at hand.<sup>10</sup>

The central dilemma for archivists is simply this: not all records having archival value can be kept. This is especially true for voluminous series of case files. While keeping all valuable records may have been feasible and even desirable for records from earlier centuries, when there were many fewer records created and the citizen-state interaction was much simpler, society would now "regard such broadness of spirit as profligacy, if not outright idiocy. Instead, archivists — like most residents of the real world — must pick and choose."<sup>11</sup> The purpose of this article is to help archivists do their picking and choosing.

### ***Definitions and the Broader Appraisal Context for Case Files***

In a world increasingly quantitative in its analytical framework, especially with regard to the computer, the use of sampling has become pervasive. Applications range from product-testing in factories to marketing studies to public opinion polling on every

conceivable subject. Sampling is based on the idea that, if a portion of the whole is properly chosen, one can safely generalize about the characteristics of the whole by studying only a small fragment of it — perhaps nineteen times out of twenty with a four per cent error rate. While this lack of guaranteed certainty may make some archivists nervous, archives have a vested interest in sampling. If archivists can with reasonable assurance make statistically valid information available about an entire body of records by examining and preserving only part of it, both archival work and scholarly research will be facilitated and scarce archival resources saved. Furthermore, because sampling is used so extensively in society and in government institutions, many records made available to archivists — especially electronic ones — are already samples. Archivists thus have an increasing need to understand better the challenges and opportunities of sampling.

Yet “sampling” *per se* is not an end in itself nor a subject best studied independently. It is only a means whereby one of several appraisal options for case files may be implemented.<sup>12</sup> Furthermore, as a term, “sampling” is employed ambiguously in archival literature, for too often it is used to refer generically to any decision to retain less than the whole population of a given phenomenon.<sup>13</sup> With this connotation, the term naturally encompasses selecting and exemplifying as well as sampling, and thus such usage is not very helpful. For that reason, working definitions of the sample, the selection, and the example are briefly discussed in the next three paragraphs, all within the context of case files. These definitions will be used consistently hereafter to reflect only the following specified meanings.

*Sample* : The choosing of items or files from a series in such a way that the items or files chosen are **a reliable representation of the whole** from which they were taken — or of a predetermined significant characteristic(s) or subset(s) of the whole. The result of “probability” sampling, as statisticians call this practice, is a statistically valid sample, the representativeness of which can be mathematically verified within an acceptable margin of error, in comparison to the original population. This means that the characteristics and relations of the whole and of the part are the same. Only this narrow definition of the statistically valid sample is a true “sample” for the purpose of archival appraisal. All other “non-probability” methods of what is often termed “sampling” are not really sampling methods at all, but rather means of selection.

*Selection* : The choosing of individual items from a series to obtain **a qualitative reflection of some predetermined significant characteristic of the whole**. There are two different types of selection. **Exemplary selection** takes from the whole some grouping of case files that the archivist judges worthy of retention, using an easily identifiable or factual criterion: all female cases in the series; all files terminating before 1966; all files appealed to a senior tribunal level in the agency. **Exceptional selection** takes from the whole the individual cases judged to have value, using some subjective criterion: the unusual, controversial, famous or precedent-setting cases, for example. Exemplary selection focuses on a collective grouping of files, not on any individual file. The intent of exceptional selection is the opposite. The result in both cases does not reflect the whole (rather it intentionally distorts it), and thus neither method has any statistical validity.

*Example* : The choosing of one or a few specimens from a case file series solely to illustrate administrative practice at the time, the forms used, the procedures followed, the levels of decision-making, the internal flow of information, and so on. There is no

attempt to be representative or reflective of the whole, and research use is usually limited to the narrow confines of strict evidential value. It is obvious that the “example” is a form of “selection,” but because of its very limited size and purely illustrative purpose it is usually and properly treated separately in archival literature.

Turning from definitions, it is very important to mention, if only briefly, two broader aspects of appraising case files. Both are outside the scope of this article, but no archivist can ignore either in appraising case files, or indeed other archival materials. The first concerns appraisal theory, to determine which series of case files actually have archival “value.” The second involves a comprehensive, strategic approach for records disposal, to facilitate more logical archival appraisal. If archivists in their institutions confront these two issues, then the modest appraisal guidelines which this article presents can be more readily applied.

Appraisal theory in North America is relatively underdeveloped.<sup>14</sup> Social models integrating structuralism (where the focus is on institutions) and functionalism (where the focus is on ideas and activities) with the citizen who interacts with both (and provides a *raison d'être* for both) have rarely been considered relevant to appraisal theory. Yet these three components interact through various recording processes (which must also be analysed) to create records, and nowhere is this more relevant than at the level of the case file. The nature of that interaction and the key factors of each are essential to determining which series of hard-copy records and which systems of electronic records may have the greatest value. While European archivists have long advocated that archivists should understand how society functions and how it creates records before appraising the actual records themselves,<sup>15</sup> North Americans have been content to search pragmatically for predefined types of “value” in records (evidential, informational, legal, etc.) and worry less about the theoretical interplay of the records creator and the functional context of records creation. However, enunciating such a theoretical model of societal dynamics that mirrors the interaction of function and structure with each other and of both with the citizen or client or customer, and the degrees of “value” that result from different types of such interaction, is **not** the purpose of this article. There is simply not the space here to repeat what has been attempted elsewhere in two companion pieces.<sup>16</sup> Instead, this article simply assumes that the archivist has already determined, through the comprehensive appraisal process outlined below, that the case file series in question have some such value, and that the archivist is now seeking ways to keep a small portion of the whole that will maintain that value.

The second broad aspect of the appraisal process concerns a strategic approach to records disposal and archival appraisal. For larger jurisdictions and institutions, this approach is designed to allow the archivist to focus on those institutions and those series and systems of case files within institutions most likely to have archival value. Records are thus appraised in their logical context rather than in isolation.

The first step in this strategic approach is to divide into priority categories the total number of records-creating institutions (and parts of institutions in the case of large agencies) for which the archives is responsible, and to obtain the approval of the institutions themselves to proceed with appraisal in a manner likely to facilitate rapid disposal and yield the highest archival return. This is a new approach for archivists. By it, they are required to “appraise” institutions and functions **likely** to create valuable records long before they appraise any records *per se*. Alas, there is again too little space

here to repeat the criteria used to rank institutions according to greater or lesser importance. Yet such criteria — functional variety, hierarchical position, size, budget, societal influence and many more — do not amount, it should be stressed, to a **theoretical** model of appraisal for determining societal interactions and archival “value,” but rather are a **strategic** or logical way of approaching the latter, given the reality that not all parts of all institutions can be appraised at the same time.<sup>17</sup>

Once the areas and programmes, and systems and series, of an institution (including cross-institutional factors) have been identified using this strategic, research-based, “macro-appraisal” approach, then the actual appraisal of the records themselves can begin, unit by unit, in order of logical priority. This second step of the strategic approach must be a “comprehensive one” — involving **all** records in **all** media within the context of a particular organizational unit or programme or function. Within this comprehensive framework, archivists should consider **first** the value of records created by the policy function and **then** those resulting from its general operations, interpretation and modifications (as revealed in policy and subject files, and in programme analysis and evaluative studies), **before** they attempt to understand and appraise correctly the records generated by the daily implementation of that policy (as revealed in the case files). Ironically, therefore, **the last thing the archivist does in appraising case files is to appraise the case files**. To do otherwise is to start at the bottom of the records pyramid with the most voluminous and repetitive records having the least value, rather than at the top with those having the greatest value. Very often, such other related records eliminate the need for the archivist to acquire any or many case files. The same result comes from a good knowledge of related published sources, internal studies and various audit, statistical or investigatory reports which may often contain aggregated, summary or extracted/sampled information from the case files. This top-down, comprehensive and contextual approach to the entire information universe will curtail excessive time delays in the disposal process, unfocused appraisal decisions, and the resulting duplication of archival records.

A corollary to this comprehensive approach is what American archivists have termed the “cluster concept.”<sup>18</sup> If there are several interrelated series of case records — military records involving individuals might include the personnel file, court martial files, hospital records, burial files, and so on — these should be appraised together following the above schema, so that overlapping information may be more readily identified and thus a better appraisal made. The same clustering occurs in immigration and certain court records, and doubtlessly elsewhere.

The rest of this article is based on the premise that the two, broader tasks mentioned above have already been performed. The archivist has some conceptual or theoretical basis for determining archival value, and has appraised in context all the relevant records in all media for a particular programme or agency or function. Having done so, the archivist now faces one or more series or systems of case files. At this point, the guidelines in this article take effect.

### *Analysis of the Characteristics of Case File Series*

Confronted with a series of case files, archivists must examine very carefully the contextual characteristics of the series as a whole and the “generic” nature of the records within each file. This research is essential in order to decide on the most appropriate

appraisal option. As will be seen, the results of this study will have a direct bearing on whether it is possible to take a sample at all, on what type of sampling methodology should be employed, on whether the sample must be taken two or more times (i.e., stratified) to account for differing variables among the files, and on whether a selection method should be used to supplement or replace the proposed sample.

Through such research, the archivist must determine whether the coverage of the programme upon which the case file series is based is universal for the entire (or at least entire adult) population of the country, city or university, or is limited primarily to special groups (armed forces, immigrants, aboriginal people), classes or occupations (miners' health files, students' applications, pilots' licences), gender (family allowances), ages (pension claims or educational grants) or regions (agricultural subsidies, fisheries allowances). As a corollary, is participation in the programme mandatory or voluntary, which will evidently affect the inclusivity of the series of records? For programmes administering grants and the like, does the coverage of the series include the rejected and unsuccessful cases as well as the accepted and successful ones? Within the coverage, are all cases available and documented in a standard way or are some consciously cut out (no file created), batched, lost or overlooked, or intentionally removed in the normal course of business to other registries and systems? Do the files collectively cover the same time span as the programme under which they were created, or just a portion of it? Similarly, do micrographic or electronic versions of the hard-copy case file exist for the entire programme, or just a portion of it? Has the documentation remained in raw, unprocessed form (microdata) or has it been replaced or supplemented by aggregated or public use data? Are exceptional, controversial and precedent-setting cases created or maintained separately from the routine and regular cases, by means of different file jackets, numbers or colours, or by indexing or abstracting the relevant information, or by creating additional files (perhaps at an appeal or special hearing level)? Do the labelling or numbering or organizational conventions used to control the series reflect an alphabetical, hierarchical, gender, ethnic, religious, geographical, regional, chronological or numerical bias — or a combination of these? Despite the apparent homogeneity of a file series's general content, format and physical appearance over time, were there significant changes in the administrative structure, agency personnel, mandates and policies, or even legislation of the programme, that may have affected the files' internal content? Were there electronic system changes or data migrations that had the same effect part-way through the programme's existence? Are there various levels in the bureaucracy (headquarters, region, field) — rather than a single level — that created documentation on the individual's interaction with the institution, whether on one central file or on several files kept in each office of the administrative hierarchy? If there are several files or several levels of bureaucracy interacting with the individual, are there formal or informal linkages of the resulting information?

In so studying case file series, archivists must ensure that they do not give undue weight to various sub-categories of case records. They cannot answer the above questions for a large series of case files by "spot-checking" or by accepting the word of the agency's officials that various records are duplicated in other series or at other levels of the administrative hierarchy. Archivists must approach the task more comprehensively and scientifically. In appraising 135,000 cubic feet of Department of Justice litigation case files in the United States, for example, archivists followed the department's own classification system to break the cases into 194 distinct categories

(kidnapping to insurance fraud) and then used a consistent methodology to select a balanced number of files from each category for study **during the appraisal process**. This is using sampling for appraisal purposes rather than for acquisition and transfer (and is sometimes referred to as “double sampling”), in order to define more precisely the nature of the information universe prior to its analysis by the archivist. As the number of Justice cases ranged from more than 10,000 in each of the anti-trust, land and taxation categories to fewer than ten for those relating to misuse of insignia, census violations or farm loans, such careful categorization is necessary in order to ensure that cases with few instances are not overlooked and those with many are not overemphasized. The Department of Justice methodology is not only directly relevant to the case file series of other judicial, court, police and intelligence agencies, but also to any series which on the surface appears to be homogeneous, but which in reality has various internal categories or functions.<sup>19</sup>

### ***Five Appraisal Options for Case Files***

After the above research has been completed into the characteristics of the records, the archivist is faced with making one of five decisions for case file series: retain all the records, extract key documents from larger files, sample or select the series, retain an example, or destroy all the records. Each option is analysed briefly below.

1. ***Retain all records permanently.*** Very few hard-copy case file series should be kept in their entirety, with the exception of certain “essential” series (see below). Those that are must have very high value and very low volume. Perhaps a small series of case files removed from a much larger programme and retained in a senior tribunal office would be an example, or a very old series of pre-Confederation records where few related records survive, or a very small series of senior scientists’ specialized research grants. As a working rule in such instances, for closely interrelated series of records, it is preferable to keep all of a small series rather than samples from a much larger one. As another working rule, the electronic rather than hard-copy version of case file information should be preserved. Where there is not a full duplication of **the archivally valuable information** between the two, the electronic version should be kept first and then supplemented by a sample or selection of the hard-copy files. Electronic records usually aggregate the more amorphous informational value and make clear the relevant demographic and statistical patterns buried in the hard-copy case files. Appraisal of records in electronic systems will include all input and output material, in every media. No small part of such input and output material finds its way into case files, thus rendering them so much less useful **if** the electronic record is identified first and protected.

Certain “essential case file series” are the exception to the rule and should always be preserved in their entirety, for their importance is incontestable to providing the core demographic profile of the nation, to protecting individuals’ legal rights, and to ensuring continuity of government administration. Examples are records proving civil status (case records of births, deaths and marriages, as well as divorce and adoption, citizenship and naturalization, and aboriginal status), land registry records, certain court and legal records (judgments, wills), and the national census of the population.<sup>20</sup>

2. ***Retain only key documents from the files.*** Immigration landing record forms or medical and employment history charts, once removed from the case files, may

render what remains behind non-archival. To remove such documents from large series is, however, a labour-intensive option if this work has not already been performed by the creating institution in the course of its normal business. It is, nevertheless, good records management practice (in which archivists have an obvious stake) to ensure that key forms can be readily separated from ephemeral material. If the forms are the only important part of the file, and they were not removed by the creating institution but easily can be, archivists should overcome their traditional phobia about file "stripping" and remove the forms. The information on the key forms (rather than the forms themselves) may also be extracted or abstracted into electronic databases, briefing books or published or near-published material, which serve the same purpose, and make the original hard-copy case files, collectively, so much less valuable.

3. *Take a sample or selection of the records.* There are several possibilities here which retain a portion of the whole, but for different purposes. As noted in the definitions section above, some accurately **represent** the entire universe, some **reflect** only certain characteristics or features of the whole, while some simply **isolate** those exceptional cases that are quite uncharacteristic of the whole. The rest of this article explores this option in much more detail.
4. *Take an example of the records.* This involves retaining a very small specimen (a file or box per year, perhaps) solely to illustrate the forms and processes involved. However, it is more sound archivally to document such evidential value of a programme by preserving, through one central point in the creating institution, the procedures and forms manuals, as well as related subject files, covering all series in an institution. Where this is possible, the "example" should be used sparingly for voluminous case file series.
5. *Destroy all the records.* This should be the decision taken for the majority of series of hard-copy case files and for many electronic databases containing case-level information. As noted before, whenever what is valuable in a series or system of case files can be documented through the other sources identified in the comprehensive appraisal process, it should be, and the actual case files themselves then destroyed. The most likely other (and far less voluminous) sources are aggregate or summary data/reports, extracted forms or decisions, subject files and publications. For hard-copy case files, an additional "other source" that renders them unarchival will often be electronic databases. Beyond this, the information in many case file series is not valuable at all in any form, being commonplace and excessively routine.

At any stage in the appraisal process, archivists can consider requesting that the institution convert the records or key information in them to electronic, micrographic or optical disk formats as an alternative to collecting extensive series of bulky paper records, or to solve conservation and perhaps privacy issues. They can also consider alienating all or some of the records from the archives of first jurisdiction to another repository.<sup>21</sup> However, these are preservation or political issues, not appraisal options. The records in such instances have already been appraised as having permanent value before the practical and preservation concerns of actual transfer and acquisition are considered. Unless such media migration has been done by the creating agency in the course of its normal business, however, archives will normally find that the conversion costs outweigh the storage ones, and only a small portion of their holdings will be so

treated. Even where microfilm or computer versions of the record are available, archivists are cautioned that such miniaturization is no substitute for sound appraisal. **Keeping useless records, even those with modest space requirements, complicates archival description unnecessarily, and clutters the desired total image of society that should be left for posterity.** Appraising case file series is no different from appraising subject files: even if archivists could keep everything, they should not do so. The role of the archivist is to preserve the clearest image possible of contemporary society and of its records creators by choosing the best records, not to add indiscriminately to the chaos of the information explosion by keeping too much or by keeping that which distorts or duplicates the image of the past.

### *Appraising the Series as a Whole: The Sampling Option*

Formal sampling is “never a ‘best’ option in appraisal, and should be considered for implementation only in the most exceptional circumstances.” Jenny Dean and Wendy Southern, two recent Australian commentators on the subject, add that sampling “should not be used as a convenient way to avoid difficult appraisal decisions relating to the research value of the records, or a means of retaining some of the records ‘in case’ they have research value.”<sup>22</sup> In sampling, some statisticians refer to the “DAM” principle — that what you want is a dam[n] good sample, which means one that is desirable, attainable, and measurable!<sup>23</sup> **Before sampling is even considered, therefore, the archivist must be certain, after following the comprehensive approach to appraisal, that a sample is desirable: that the information in the case files is substantial in nature, that it is needed for quantitative research into the collective character of the series, and that there is no other possible way of obtaining this or similar information from other sources.** It is thus not always desirable to keep a sample, even if one is attainable and measurable. And even if a sample is desirable, it may not be attainable or measurable unless certain preconditions, as noted below, are present to ensure that the results will be useful to researchers. More than any other appraisal option, sampling is demanding, scientifically precise and time-consuming, and therefore should only be used where there is no alternative.<sup>24</sup>

These cautionary notes concerning sampling are based on the fact that its application in archival appraisal is not straightforward. Due to the irregular circumstances which usually surround the creation, management and disposal of records, the preconditions required to ensure the statistical validity of formal or “probability” sampling often cannot be met in an archival setting, as will be seen shortly. Yet the use of “non-probability” sampling methods by archivists, which unlike “probability” methods do not have statistical validity, cannot be recommended. What may be appropriate for product-testing or some kinds of opinion-polling is not so for archival records chosen to reflect the whole with statistical validity, for all time, for all manner of research uses, in all types of disciplines. A brief discussion should demonstrate the truth of this assertion.

Non-probability sampling takes a sample of whatever is at hand and extrapolates from it general conclusions relating to the whole. There are two types of non-probability sampling: “convenience” and “purposive” sampling. Convenience sampling is often used in commercial or news-gathering situations. Examples might be surveying the first 100 voters to leave a polling station, checking the expectations of the next twenty-five guests to enter a remodelled hotel, or assessing the opinions of 200 people on new pickles in a grocery store. As may be imagined, similarly retaining the first 200 files as a

“convenience sample” of a large case file series in an archival setting would be next to meaningless. For a large national programme, the result might well include only those people living in Newfoundland whose surnames start with “A” or “B.” Purposive sampling **appears** to be more scientific, for by this method experts try through research to isolate a characteristic or feature of the population that they judge to be representative of the whole, and then choose a sample based on that characteristic with the intention of reflecting the whole.<sup>25</sup>

The dangers of non-probability purposive sampling in an archival context can be illustrated by the “F” sample.<sup>26</sup> Based on an extensive analysis of surnames correlated to Canadian ethnic, linguistic and geographical factors, the “F” sample (choosing surnames beginning in “F” to yield a 3.54 per cent representative “sample” of the entire Canadian population) enjoyed some currency at the National Archives of Canada in the 1980s. It is no longer used there and the reasons for this change demonstrate, by implication, the problems of using “non-probability” sampling methods in archival settings. In the first place, there was no statistical validity to the “F” sample, because one of the key requirements (as will be seen) of formal sampling is that all members of the population have an equal “probability” (or opportunity) to be chosen, which with the “F” is evidently not the case. Furthermore, as the “F” was derived from a cross-Canada analysis of surnames for the 1950s to the 1970s, it is fixed in time. To maintain its validity now, and in the future, the National Archives would be faced with regular and expensive updating of the surname research data and tables. For series running before or after those dates, for those series not representative of the Canadian population as a whole (such as the armed forces, aboriginal people or Quebec-based programmes), and for the vast majority of regionally-based case file series, the “F” is simply invalid. The ethnic mix of the population represented in local series of records in Vancouver, Montreal and Halifax, for example, is in a far different proportion from that of the Canadian population as a whole. More seriously, the “F” rarely occurs in East Indian and ethnic Chinese surnames, and the new romanizing conventions for Chinese characters have removed those few to whom it did apply: the Fongs are now Honggs, Xuonggs, and other variations, while the Indonesian Fo is now Pho. Prefaces in German and French especially — de, du, von, y, etc. — mask many “F” names and there is simply no “F” at all in Korean. Finally, the “F” analysis did not include economic or class variables, which are central to the character of many series of case files. Thus, despite the best intentions in the world, and extensive expert research into population characteristics, the “F” method is seriously flawed.<sup>27</sup> It demonstrates that non-probability sampling methods are rarely appropriate for archival sampling, if the aim is a sample that will be an accurate, statistically valid representation of the whole.

Turning to probability sampling, which is the only kind of sampling recommended to archivists in this article, there are several required preconditions. To achieve statistical validity, it is highly desirable that a complete set of the records representing the entire population of the programme be available at the time the sample is taken. In other words, as formal sampling requires that all members of the population have an equal probability to be selected, the total population must be fixed and known.<sup>28</sup> To ensure a complete set, the records must also be accurately described and numbered. Furthermore, the sampling method must be based on random selection. One statistician points out “that in all types of probability sampling there must be both some element of randomization and some sort of complete listing . . . It will always be necessary for the

researcher to examine his list carefully and to know how it has been constructed and the nature of its defects.”<sup>29</sup> Except with great difficulty (see below), one statistically valid sample of the whole cannot be taken from an incomplete or open-ended body of records, or from a body of records which is inaccurately described, or from one where the exact population is unknown or unknowable. Sampling is, therefore, most appropriate for closed series, which are tightly controlled, numbered and finite.

Case files suitable for sampling must also contain a very high degree of subject and document homogeneity and a low degree of individual file or document variation. Sampling is most appropriate, therefore, when used for particular instance case files or forms, as well as for electronic records, which usually relate to a single transaction or event, usually concluded over a short time-span (vocational training, immigration entry or mortgage loan application files, for example). Increasing variation (lack of homogeneity) enters with continuing events case files, where numerous events over a longer period of time are recorded on a single individual or organization (medical history, student, criminal or employee records, for example); or with programmes that encompass many free-form letters and memoranda rather than standard forms and questionnaires and model paragraphs; or with extensive regionalization and fragmentation in a programme’s delivery; or with appeals, hearings, and other disputes transforming a usually simple particular instance case file into a complex one (the immigration deportation process, for example). Such variation, or relative lack of homogeneity, means that the sample must be larger, or stratified and taken several times over, in order that all the significant characteristics of the entire population are represented. Again, as noted earlier, this underlines the necessity of the archivist studying carefully the nature of the series and file contents and the registry system organization. Sampling error most often relates to this failure to detect significant variations within the records, thus resulting in a sample that is not representative of the whole. If detected beforehand, such variation can be compensated for through stratification. If not, then the statistical variance of the sample will most likely be unacceptable (that is to say, mathematically, the values of the sample and the whole population will vary by too wide a percentage for the sample to have statistical validity). In lay terms, the sample will not be an accurate reflection of the universe from which it is drawn.

The statistical validity of a sample depends on the size of the sample, as well as its being randomly chosen. The precision of the statistical inferences to be made from the sample increases with the size of the sample. This advantage, however, can be obtained without an enormous increase in the sample size or a requirement to take a huge sample. This is especially so when sampling very large populations. In other words, a random sample of 1,000 case files from a records system for a programme which created 6,000 of them in total would not necessarily be more precise than a random sample of 1,000 out of 2,000,000. In opinion surveys, as few as 1,400 people can be included in a valid sample of about 55,000,000 people. As one statistician noted, “. . . the reliability of the sample as regards the information it reveals about its parent source is hardly affected at all by the proportion the sample bears to its parent group . . .”<sup>30</sup> It is reasonable to conclude that, since the size of the populations of records which most archives are likely to be sampling in any one sampling exercise will be considerably less than, say, 55,000,000 files, sample sizes of from 1,000 to 1,400 files are probably as large as will ever be required **for a complete series**. One statistician makes a general point about sample sizes in probability

sampling which, though offered to social scientists, is helpful to archivists: "In most survey problems the research aims are many and indefinite; hence, only vague ideas surround the precision required from the sample results, but the available funds may be fixed within rigid bounds. Then the sample design should aim to obtain maximum precision . . . for the fixed allowed cost."<sup>31</sup> Since archivists also have difficulty predicting the many possible uses of samples of records, they ought to aim for as high a degree of precision (or as large a sample size) as possible, to a usual limit of 1,400 files. Much smaller sample sizes are reasonable when the population of records is smaller. It is recommended that, for such populations, the table developed by Bell Telephone and widely adopted in industry be used by archivists; it is reproduced in the Appendix of this article and, for populations ranging from fifty to 150,000 cases, sets sample sizes ranging, respectively, from five to 800 cases. Generally speaking, samples should **not** be chosen on a percentage basis: the traditional five or ten per cent sample really means nothing, and may give a result far too large or too small.<sup>32</sup>

As contrasted with the closed or finite series, sampling can only be applied with great difficulty to the continuing or open-ended case file series of the type most often faced by archivists responsible for modern corporate entities. As such series are open-ended at the time of appraisal, the information universe is unknowable, for thousands and likely millions of files are yet to be created at the time when an equal number are ready for disposal. The files in such series are usually also physically scattered in field offices across the country, as well as loosely and inconsistently numbered and controlled. It may therefore be virtually impossible in practice (if not in theory) to determine how many records belong to the series (i.e., the total sampling universe), whether a complete set of the records exists, if they have the high level of homogeneity essential for statistically valid sampling, and if the sampling method chosen can be properly applied in those offices. These uncertainties are sufficient to undermine the statistical validity of any resulting sample, and can only be removed by very labour-intensive work, as detailed below. If these uncertainties cannot be removed, no sample should be attempted, for the result would be statistically invalid and therefore lead researchers to a distorted view of society.

To address this dilemma, the archivist can artificially close open-ended or continuing series: sample the files for 1975-1980 now, the files for 1981-1985 in five years, the files for 1986-1990 in ten years, and so on. This may be possible in some cases, but it assumes a rigid consistency of file closure and dating practices rarely seen in institutions, and it means the ultimate size of the total sample remains unknown. As well, using this example, there is no guarantee that the three separate, statistically valid samples of the three separate dated portions of a series will add up to the same as one sample taken of the eventual whole series when it would have been closed. The three portions would always have to be maintained and used in the archives as three separate groupings, with researchers applying appropriate weighting formulae when they use combinations of the files bridging these groupings. The problem of this fractured or multiple sampling is apparent. In an expanding series, to continue the above example, a sample of 1,400 cases might be chosen out of 140,000 in 1991, a sample of 1,400 out of 280,000 in 1996, and a sample of 1,400 out of 1,400,000 cases in 2001. While each sample alone would be statistically valid, the archives could not mix the three samples, for in the three samples, each case had, respectively, a one in 100, 200 or 1,000 probability of being chosen. Therefore, if the three samples of 1,400 each were later merged at the archives into one

sample of 4,200 files, the result would deny the first requirement of statistical validity: that each case in the population has an equal probability of being selected, which in this — and in almost all imaginable similar cases — it did not. And this example is relatively straightforward. Imagine in just one agency thirteen important programmes (out of more than fifty), each generating case files, in 1,000 local offices, where records because of high volume and space shortages are disposed of annually! After ten years, the archives would have 130,000 samples which, to maintain their statistical validity, could not be merged intellectually (even if some could be physically), and this simply for one fonds! The example is not hypothetical, but taken from the appraisal of the records of the Canadian Job Strategy-Employment Services sector of Employment and Immigration Canada.

The problem is easier for those few continuing or open-ended case file series which are unified, sequentially numbered, and maintained at headquarters or some other central point. Here the records manager can be instructed to pull for the archives on a continuing basis every 500th or 1,000th or *n*th file from now until the series ends. This will gradually build up a sample at the archives, but it requires excellent numbering and control of the case files, over time, by the institution, and if the records creation rate alters significantly control over the eventual size of the sample will be lost.

It is also possible to use projection analysis to cope with the unknown information universe. Such analysis uses known populations and growth-rate factors to predict the size of future populations. This obviously requires great mathematical sophistication, and the result is quite uncertain. It is not recommended as practical in archival settings.

Another alternative is to delay the taking of any sample until the series is closed (this must not be considered for those electronic records where data are regularly deleted from a system and therefore lost to the archivist). Given the very high volumes which case file series often involve, the very short retention periods in most instances of individual files in the series, and the sometimes extended operation of their parent programmes over many years, such a strategy would involve, if applied to all major government institutions, storing literally hundreds of millions of case files in the records centres or on selective retention at the archives pending a decision. **Except in the rarest of cases, this is not a responsible use of resources, and a difficult decision would have to be taken either to sample the series in stages or, declaring a sample unworkable, to destroy it entirely.** This is another general rule of sampling: if a sample is the only archival choice that makes sense, and it is impossible to take that sample, then all the records are to be destroyed. As Dean and Southern observed, a sample is not an excuse to take “something” in order to assuage an archivist’s guilt, but a sensitive procedure that must be implemented precisely or not used at all.

Therefore, for continuing or open-ended high-volume series of case files, the collective archival value of the case files must be extremely high to attempt sampling at all. Sampling of such open-ended series should be limited to single headquarters applications, or in special circumstances to major and self-contained regional offices, and in both cases only to where there is evidence of excellent, centralized records numbering, control and disposal. Even so, in such circumstances, control over the size of the final sample will be lost. Samples of continuing or open-ended high-volume series of case files should never be taken at the local or field office level.

There are numerous methods for sampling. To summarize the foregoing points in one sentence, all require randomness, which may be obtained by using random number

tables or an automated random number generator; all require members of the population to be uniquely (preferably sequentially) numbered and counted; and all require that every member of the population has an equal probability of being chosen. Three common types of sampling are simple random, systematic random and stratified random. **Simple random sampling** is where the 1,400 chosen numbers are applied randomly across the entire population. It is thus most appropriate for very homogenous, very short-lived series without significant internal geographical, gender, chronological or other bias. If not used in such limited circumstances, strange results can occur. For example, in a series with 1,000,000 files, it is feasible that 75 per cent of the randomly chosen numbers might range between one and 50,000, rather than be spread evenly over the whole 1,000,000 numbers/cases. This means that pockets of files of a particular type may be missed entirely, depending on how the series is organized and classified. If, for example, the 1,000,000 files were organized using a west-to-east geographical file number code for some national programme, then this particular sample would contain 75 per cent of its cases from British Columbia, rather than the approximately 12 per cent representation that that province should have. Similarly, if the same series were organized, as is often the case, in chronological order, where each new case gets the next sequential file number, then 75 per cent of the sample would relate to only the first two years of a forty-year programme.

Perhaps the most appropriate method for archivists, therefore, is **systematic random sampling**, where only the first number is chosen randomly, and then every *n*th number thereafter is chosen until the full sample size (say 1,400 cases) is attained. This method avoids the “missing pockets” syndrome of simple random sampling. It may be helpful to explain the methodology of systematic random sampling more carefully, as archivists will probably use it most often. If there are 700,000 files in a series, and a total sample of 1,400 is required, the archivist will be taking  $(700,000 \div 1,400)$  every 500th file. (This assumes, of course, that the files are organized more or less chronologically and that there are no formal numerical subdivisions by region, occupation, programme, case type or other “strata.” If so, then stratified random sampling would be appropriate, which will be explained below.) Therefore, using random number tables or a generator, since every 500th file will be chosen, the numbers from one to 500 should be used/entered in order to choose one number at random. If the number 376 came up, then the sample to be pulled would include files numbered 376, 876, 1376, 1876, 2376, and so on until all 1,400 files had been chosen. This method could be combined with a time-series method: if 700,000 files were created annually, for example, and an annual sample of 1,400 files was considered excessive, the archivist could perform (like the census) the sample only in years ending in “1” and/or “6” instead of every year. Where the entire universe is present and organized by social insurance number (SIN), the use of SIN terminal digit 5 or 55 or 555 (for a 10, 1, or 0.1 per cent sample, respectively) is another type of systematic random sample. Such a method could be used for moderate-sized series after the mid-1960s; for very large series, the resulting sample would be far too large and too difficult to pull.<sup>33</sup>

**Stratified random sampling** is where the whole is broken down into logical “strata” (which may be defined as parts or subgroups or geographic areas or file blocks of the whole — like the categories in the United States Justice Department litigation case files mentioned earlier), and then each stratum is randomly or systematically sampled, thus ensuring that no part is overlooked.<sup>34</sup> Stratified sampling may also be used to acquire two or more samples from the same series in order to protect different characteristics of

the whole, especially when the archivist believes that the participation of a particular group (or groups) in a programme was significant. This could include lower courts and appeal courts, different income levels, male and female subjects, different age groups, different regions, different linguistic or ethnic groups, or any other strata or subsets into which it would be useful to divide the files. An example might be where a random sample of about 1,400 files was taken from a series relating to a nation-wide investment programme. Then, because aboriginals made up only a tiny percentage of the population of this particular programme and might therefore have been missed by the first sample, or at least represented in insignificant numbers, and since the archivist judges aboriginal participation in this particular entrepreneurial activity to be very important, a second random sample of 100 or 200 cases which dealt only with aboriginals might be taken from the remaining case files. This would provide the researcher with an obviously weighted sample, but would allow for analyses comparing aboriginal and non-aboriginal participation in the programme.<sup>35</sup>

The advantages of sampling may be simply stated. Unlike any other method, the result can be used to reconstruct the whole with statistical validity.<sup>36</sup> It thus facilitates accurate quantitative research for a multitude of disciplines and interests. Sampling is theoretically unbiased and thus easily explainable to researchers. For a numerical arrangement of files located and controlled centrally, it is relatively easy for clerical staff to pull. Finally, archivists can control the size of the sample. Normally, that size will be quite manageable, since even for large series, the proper statistical weight can be assigned even when a relatively small sample is chosen.

However, there are also certain limitations to archival sampling, in addition to the issues already mentioned concerning the open-ended series and the regional dispersion of records. There is obviously little chance that the few exceptional or outstanding cases in the series will be included in the random sample, although this can be compensated for by using a selection method (see below) to complement the sample. As well, researchers cannot do longitudinal work: it will be impossible to trace a particular individual or office or county over time, as the county or person or office in all likelihood will not be selected for every annual or decennial random sample from the series. (This is different from time-series sampling, like the census itself.) For files arranged alphabetically or in some other non-numerical scheme, the statistical sample is very difficult to pull physically, as it will require the counting — and often may require the costly numbering — of all the files before pulling. And for complex file series, there may be the need for a stratified (i.e., multiple) sample to ensure that various types of actions are sampled; this is very expensive and requires statistical and analytical expertise, especially where the strata are not obvious in the way the files are classified, numbered and labelled.<sup>37</sup> For stratified sampling especially, a high level of analysis is also required to determine the homogeneity of the series and the nature of the features or characteristics within the files which must each be given statistical weight. Archivists naturally should not be afraid of complex analysis or of acquiring new expertise, but only cautious that the time thus spent to determine these factors does not pass the point of diminishing returns. **Sampling is a powerful tool, therefore, but should be used sparingly and only when all the conditions for statistical validity can be met and all other appraisal options have been considered first.**<sup>38</sup>

#### *Appraising the Series as a Whole: The Exemplary Selection Option*

Exemplary selection chooses groups of similar files from a series to obtain a **qualitative reflection** of the whole or of some predetermined significant characteristic of the whole.

As with sampling, the focus is on the collective grouping of files chosen, not on any individual file. Where a sample is impossible or impractical, or more often unnecessary given the existence of other sources (electronic records, aggregate reports, etc.), the exemplary selection method can be used to gain a sense of the qualitative information “typical” of the series or some important feature of it. Such information often provides the “human dimension,” the “local colour,” or the “quotable quotes” used by researchers to supplement the information found in databases or summary reports and statistics. Exemplary selection uses an **easily identifiable or factual criterion**. Examples might include pulling all Chinese immigrant cases in a series; all files terminating before 1940; all files containing a “notice of appeal” form; all files from the years immediately before and after an agency reorganization or significant legislation to show the **collective** impact of such changes on actual operations; all files for particular types of court proceedings (e.g., felony convictions, kidnappings); or all personnel history files for military figures reaching the rank of “Major” or above. A different type of example might involve choosing every 150th box or every fortieth file in a large or small series, respectively, with no element of randomness involved and no assurance of the nature or number of the total population (otherwise such a method would be a sample). Taking another approach, the archivist could keep all series for an agency for a very intensive geographical area or areas (a small region or city or office) which, by itself or in combination with a few others, is “typical” of the whole nation, in order to take a high-quality snapshot of the societal image. Finally, all files bearing some significant physical characteristic might be chosen, such as all “fat files” (to which concept I shall return in more detail) or all “secret” files or all sensitive files in cautionary red folders — if this physical characteristic is judged by the archivist to be the important variable distinguishing valuable files in the series from those without value. All these factors will naturally vary from series to series.

The advantage of the exemplary selection is that it can be used to trace, at least impressionistically, the operations of a programme over time and to provide qualitative colour. It may be a reasonable compromise where sampling is impossible: keeping all of several open-ended series for forty-five Canada Employment Centres chosen for their geographical representativeness is preferable to trying to implement a sampling scheme for each series in each of a thousand such offices. Moreover, where a complete collective subset has been selected from the whole — all the women prisoners or Chinese immigrants, for example — statistically valid research can be done on the subset. Because the exemplary selection is based on an objective or physical criterion, it is easy for records clerks in departments or records centres to pull the required files from the whole.

The method certainly has limitations. It is not statistically valid and cannot be used to reconstruct the whole or to do any quantitative research relating to the whole. With some **partial** exceptions — flagging the appeal notice files, the hierarchical cut-off (all Majors and above), the “fat file” method — it does not save **individual** exceptional cases, but only **groupings** likely to be exceptional. A Nobel Prize-winning author may well not have achieved the rank of “Major” in the army; a very famous immigrant may well not have generated a “fat file.” This underlines that the result is collective, not individual. Moreover, there is no control over the eventual size of the selection, although this may be estimated in advance and adjustments made accordingly. It does require research expertise to make the right choices, as the “typicality” of the isolated feature,

characteristic, time period or geographical area will always be open to debate, and therefore will require the archivist's careful analysis. In this way, it is no different from the problems confronting non-probability sampling (such as the "F" sample), although unlike that approach, exemplary selection makes no claim to statistical validity.

### *Appraising Special Cases Within a Series: Exceptional Selection*

Exceptional selection focuses on special cases within the case file series worthy of archival retention. If the series as a whole does not have any collective value as outlined in the previous two sections, then this final step takes place by itself. If the series as a whole does have collective value, that must be determined first, before the search for outstanding individual files or items within a series commences. Unlike the previous two methods, exceptional selection rarely applies to electronic records.

The National Archives and Records Administration (NARA) in the United States has outlined specific criteria for identifying special cases most likely to have archival value at the level of the individual file. These include any file which established a precedent and therefore resulted in a major policy or procedural change; was involved in extensive litigation; received widespread attention from the news media; was widely recognized for its uniqueness by established authorities outside the government; or was reviewed at length in the agency's annual report.<sup>39</sup> These criteria have been applied successfully by NARA to series involving research grants awarded for studies; research and development projects; investigative, enforcement and litigation case files; social service and welfare case files; labour relations case files; case files related to the development of natural resources and the preservation of historic sites; public works case files; and court case files.

In addition to such **exceptional** files in their relation to the programme of the agency, some archives may wish to keep **routine** case files on noteworthy people for the biographical information such files often contain. The military service file of someone who later became prime minister or the immigration file on a later terrorist might be examples. However, extreme caution must be exercised here. Most archives can no more be the repository of such random biographical information than they can of genealogical information at the level of the ordinary case file, **or else by definition every such file would have to be kept**. Except where old series of such files survive almost by accident, however, this may be a non-issue. Because of the short administrative life and very high volume of such case files, they usually have very short retention periods, and therefore are destroyed long before their subjects become "famous."

Unless the creating agency is willing **during the normal course of business** either to code the file jackets (numbering or lettering variation, colour tabs, a cover stamp or hand-written annotation), or to separate them physically into special categories in order to indicate that a particular case file was indeed exceptional and falls into the categories enunciated by NARA, or relates to a "famous" person, there is little chance that the archivist will be able to isolate such files using these categories — especially if there are hundreds of thousands or millions of file units being disposed of regularly. A second method is to require the operational officials of the institution — not the records management or records centre staff — to review all the case files at the time of disposal to identify such important individual cases. Whether institutions would be willing to do this or, if they were, whether years after the case was closed the expertise remained within the institution to identify many such files, remains highly doubtful.

Because of these difficulties, the archivist usually has to adopt other, less satisfactory, means to identify and separate important cases from a series. This involves isolating groups or categories of records **likely** to contain the most unusual or controversial individual files. On the surface, this appears to be the same as exemplary selection, but the purpose is quite different. In exemplary selection, the aim is to isolate some typical collective grouping to mirror the quality of the whole or some aspect of the whole; with exceptional selection, the aim is to use the grouping only to isolate indirectly the unusual individual files that are not typical of the whole. Four such collective or group approaches may be recommended:

1. *Isolate important cases by date*: military records during wartime years; immigration records during years of special migrations or forced evacuations, whether globally or by particular countries; all files created during the pioneering or controversial periods of a particular programme. Of course, for some programmes, this could amount to every year, but for others, such a time-based method might be a useful tool to separate the more interesting cases from the routine.
2. *Focus on certain levels or categories of individuals*, where such hierarchical organization exists and is easily evident in the filing system used for the related records, and where such upper levels coincide with the particular significance the archivist wishes to retain concerning the programme. Public service civilian personnel records are a good example, where all files at the federal level, for example, are preserved of persons reaching the rank of Director (or equivalent) during their careers.
3. *Concentrate on those areas of institutions (or related outside agencies) where the unusual and controversial cases are handled (or referred or adjudicated) as a normal part of daily operations*. Tribunals, appeal boards, ministerial review committees and certain courts will in this manner already contain the "problem" cases: the unusual, the precedent-setting, the controversial.<sup>40</sup> A recent example concerns the records of the federal Pension Appeal Board (PAB), where controversial Canada Pension Plan (CPP) cases are appealed. Here, about 4,000 files over twenty years have been extracted from the many millions of CPP files in the Income Security Programs Branch of National Health and Welfare. The full CPP series could never be sifted file-by-file by archivists searching for the unusual and controversial cases; in the PAB records, however, they found that this had already been done for them by the system itself.
4. *Concentrate on the "fat file"* — or, as it is more elegantly called, the multi-section or multi-volume file. As exceptional, unusual or controversial cases almost by definition generate more correspondence than their routine counterparts, such files will be thick and thus easily identifiable, even in vast series to be pulled for archival retention. The NARA Department of Justice litigation case file appraisal studied "fat files" very carefully.<sup>41</sup> NARA archivists compared fat files to thin files, and judged both against a random sample of the entire population. They assessed the contents of the fat, thin and sample groupings file by file in terms of high, medium, low and no archival value, as well as including factors concerning the variety of correspondence, level and significance of decision-making, gravity of case or offence and number of offices involved. The results are a conclusive

demonstration of the value of the “fat file” approach in isolating the most important files, according to the above criteria:

Archival Value Rating (%)	High	Medium	Low	None
Regular Sample Files	0.2	0.2	15.2	82.2
Single-Section Files	0.0	1.4	14.3	84.1
Multi-Section Files	5.1	19.3	36.2	39.2

Of course, **not all thick files necessarily follow this pattern**: it may be that someone was routinely repaying a loan in monthly payments over thirty years (thus generating a fat file of 360 receipts). **The archivist will have to assess the functional and operational reasons for the thickness of particular files in each series** where they occur, to ensure that such files are indeed exceptional to the series as a whole. What is thick for one institution or series may also be thin in another, and so the “fatness” determination must be made series by series. It is logical, however, that in many cases fat files may well contain all that the archivist feels is necessary to document the process involved in the citizen’s interaction with the institution, for almost by definition the controversial and precedent-setting cases should best reflect challenges to the agency’s intentions and objectives. The “fat file” method is also particularly useful for documenting collective evidential value, for the NARA study also reveals that more administrative processes and more varieties of institutional activity will **usually** be reflected in fat files than in their thinner counterparts.

The advantage of exceptional selection is that it saves the files usually of greatest interest to researchers who are not undertaking collective quantitative research. The limitations of this method are equally obvious: it has no statistical validity, and will always give a false impression of what the original complete series was like (i.e., it will distort the view of a “typical” case). Great substantive expertise on the part of the archivist as well as very clear prior identification and arrangement of files by the creating institution is also required so that the exceptional cases can be located. Again, the size of the eventual selection cannot be controlled, although with trial runs it may be estimated.

In conclusion, it must be noted that the archivist can also combine two or more of the above methods, where appropriate. **If sampling is one of the methods, it must be applied first so that the statistical validity of the whole is not impaired.** It may be most desirable to use systematic or stratified sampling first, and then to apply an exceptional selection method to isolate noteworthy individual cases. Of course, some of these may already have been captured first in the sample; if so, they must not be removed from the sample or its statistical validity will be undermined, but rather must be cross-referenced in archival descriptive tools.

### ***Sampling Electronic Records***

As should be clear from the comprehensive approach to appraisal recommended earlier, hard-copy and electronic case files must be considered together, especially when they overlap or deal with the same programme. Indeed, primacy should be given to electronic records as the best way to retain archival information on individuals and organizations which interact with the corporate institutions of government, business, universities and so on. While the foregoing sections of this article have been based on the assumption of cross-media appraisal, some special features concerning electronic records are the focus of this section.<sup>42</sup>

Creating electronic samples often occurs in institutions in the normal course of their business activities well before the records are made available to archives. At the federal level, in addition to the national census itself, the various surveys and questionnaires carried out by many departments, which are recorded, tabulated and analysed electronically, are obvious instances of electronic records that are first created as samples. Equally important, but often overlooked, is the fact that departments very often employ great statistical expertise to create from much larger databases their own electronic samples for analytical and policy purposes. Rather than take the time and expense to “run” hundreds — sometimes thousands — of magnetic tapes of all the microdata for huge programmes, such as family allowance or vocational training, the department will download from these large databases its own electronic samples, and/or build its own electronic longitudinal data files to manipulate for policy and programme evaluation. Not only do such samples have very important evidential value on their own, but they may well be all the electronic case records an archives should acquire — and can afford to acquire — for that particular “series.”

Archives may soon follow the lead of government departments, the sampling of electronic records by archives becoming a growing activity in the years ahead. With large databases, it is not possible, even for major archives, to acquire all the electronic records of a valuable programme. For example, the rough cost **just to supply the blank reels of magnetic tape** to acquire the master files of the family allowance programme since 1944 (about five thousand tapes), including making a back-up conservation copy of each, would be around \$200,000 — to say nothing of processing time and salary dollars. Fortunately, electronic records lend themselves to sampling far better than do hard-copy paper case files. As each case or “record” within an electronic data file automatically carries a unique “record number” and the whole universe of the “series” can easily be counted and identified at any point in time, sampling is usually quite possible. Such samples (if taken of an annual master or history file, for example) should be consistent year after year, to allow for longitudinal and comparative research. With the increasing capacity of data compaction on storage devices, however, most small or mid-sized databases (one magnetic tape in total) should be acquired without sampling, for sampling often destroys the possibility of data linkages between different data files, which is one of the major advantages of acquiring any electronic record.<sup>43</sup>

### ***Sampling and Archival Description***

Once case files have been appraised and acquired, they must be arranged and described. There is, indeed, a significant carry-over of the information required to make an informed appraisal to the information needed for a proper description. For this reason, a few guidelines on the subject may be appropriate.

Two or more *samples* from the same series of case files must always be arranged and described in two or more **separate** sub-series of the inventory and related finding aids. This relates both to samples taken in a stratified approach from a single series (one sample of the entire population, and then another of a special or smaller characteristic of it), or to a consecutive number of samples taken from the same continuing or open-ended series over many years. Similarly, two or more **different selections** (for example, all files before 1945 and all fat files) must also be arranged and described separately. Different iterations of the **same selection** method (fat files received in 1990, 1991, 1992, etc.) can, however, be arranged and described together. It is crucial to remember that

while two or more groups of case files from the same (or even closely related) series may well look the same and be numbered or labelled in the same manner, they must never be mixed together intellectually (nor preferably physically), if they were appraised and acquired for different reasons. To mix **any** of these records will destroy the statistical validity of the sample, and undermine the archival integrity of the rest. Even if there is some overlap — for example, where a few of the controversial or famous cases fell into the sample taken first, before the exceptional selection — the solution must rest in cross-references in finding aids rather than in removing anything from or adding anything to a sample.

The reason for choosing each separate part from a case file series — to sample, to isolate famous cases, to reflect geographical qualitative information — must be made clear in the descriptive paragraphs of each inventory entry. As well, for case files, researchers should be told explicitly in the inventory description whether the series involved has statistical validity or not. All the details of any sampling method used either to appraise or to acquire records must also be completely explained, including the size of the original universe, the type and method of randomness employed to create the sample, the size of the sample, the analysis of the file variables or characteristics, the reasons for and nature of stratification and any special weighting or over-sampling. Finally, for series of case files sampled or selected for their **collective** value (as contrasted to those containing famous or controversial **individual** files), the records are to be used collectively and therefore no nominal listing should be created, let alone made available to the public. In both cases, the privacy of the sampled or selected individuals must not be compromised, and all relevant privacy regulations of the jurisdiction should be observed.

### ***Conclusion***

The appraisal of case files challenges archivists to hone their professional skills. In this, sampling should not be viewed as excessively technical or arcane, but rather as one of several useful tools to implement appraisal decisions for case files. Such appraisal itself is undoubtedly complex, with many variables to consider, and it will require both a deeper and broader commitment by archivists to contextual research and a more comprehensive and strategic approach to acquisition than has often been the case in the past. Yet the rewards for doing this job well are considerable. Many case files are created, but few should be chosen. It is hoped that this article will help archivists to identify the few gems that have value and destroy the many which block their light from view.

## APPENDIX

### The Bell Telephone MIL-STD 105D Sampling Plan\*

Engineers at the Bell Telephone Laboratories devised a sampling plan for the United States Government in 1942 which, after four revisions (the last in 1963), has become an industry standard. Indeed, the International Organization for Standardization adopted it in 1973. It allows smaller samples to be chosen than the usual limit of 1,400 cases where the overall population is smaller.

#### TABLE FOR DETERMINING SAMPLING SIZE\*

**Population (Total number of items to be sampled)**

**Sample Size**

(based on homogeneity and value:  
see note below)

	<i>Low</i>	<i>Average</i>	<i>High</i>
51-90	5	13	20
91-150	8	20	32
151-280	13	32	50
281-500	20	50	80
501-1,200	32	80	125
1,201-3,200	50	125	200
3,201-10,000	80	200	315
10,001-35,000	125	315	500
35,001-150,000	200	500	800

**Note:** The "low"-sized samples are based on the least evidence of substantive informational content and the greatest homogeneity of the files, whereas the "high"-sized samples are based on the reverse: higher substantive value and greater variation of internal content (i.e., lower homogeneity).

\* The method is cited, with the chart, in Joseph Carvalho, "Archival Application of Mathematical Sampling Techniques," *Records Management Quarterly* 18 (January 1984), p. 63. This reference was kindly brought to my attention by Rod Young, who has himself done extensive reading on sampling issues.

#### Notes

- 1 This article is based in part on two larger studies I have recently written: "The Appraisal of Case Files: Sampling and Selection Guidelines for the Government Archives Division, National Archives of Canada," internal report (January 1991), which itself was informed by *The Archival Appraisal of Records Containing Personal Information: A RAMP Study With Guidelines*, International Council on Archives (Paris, [forthcoming] 1991). In completing those works, I had the advice of numerous colleagues, named therein, and their suggestions are reflected again in this essay. While I want here to acknowledge collectively my appreciation of their past support, two must be mentioned by name. Trudy Peterson of the National Archives and Records Administration gave me for the RAMP study her schema on the advantages and disadvantages of various sampling methods, and Tom Nesmith of the University of Manitoba, before he left the

- National Archives of Canada, completed useful work on the statistical basis of archival sampling. In addition, this article was read by Ed Dahl, Eldon Frost, Tina Lloyd, and Dan Moore, all of the National Archives of Canada, and by Robert Hayward of the Treasury Board Secretariat, Government of Canada. Their helpful suggestions on substance and, in the case of Dahl and Hayward especially, extensive editorial corrections have very much improved the piece, saved its author public embarrassment, and spared the reader the tedium of a much longer first draft.
- 2 The pioneering and still the best statement is Tom Nesmith, "Archives from the Bottom Up: Social History and Archival Scholarship," *Archivaria* 14 (Summer 1982), pp. 5-26. This thematic issue of *Archivaria*, of which Nesmith was guest editor, was entitled "Archives and Social History" and contains numerous articles either suggesting or demonstrating the imaginative use of case records by scholars in several disciplines to gain fresh insights into society. Two other important studies aimed at archivists are Joy Parr, "Case Records as Sources for Social History," *Archivaria* 4 (Summer 1977), pp. 122-36; and Peter Gillis, "The Case File: Problems of Acquisition and Access from the Federal Perspective," *Archivaria* 6 (Summer 1978), pp. 32-39. Beyond these more generic studies, there is a growing number of articles on the value and use of particular types of personal case records in an archival context; see, for example, R. Joseph Anderson, "Public Welfare Case Records: A Study of Archival Practices," *American Archivist* 43 (Spring 1980), pp. 169-79; David J. Klaassen, "Achieving Balanced Documentation: Social Services from a Consumer Perspective," *The Midwestern Archivist* 11 (1986), pp. 112-24, and especially pp. 118-19; and John C. Rumm, "Working Through the Records: Using Business Records to Study Workers and the Management of Labour," *Archivaria* 27 (Winter 1988-89), pp. 67-96. For another perspective on uses of such archival records, see Michel Duchein, *Obstacles to the Access, Use and Transfer of Information from Archives: A RAMP Study* (Paris, 1983), pp. 8-9.
  - 3 For only one such example, see Judith Roberts-Moore, "The Office of the Custodian of Enemy Property: An Overview of the Office and Its Records, 1920-1952," *Archivaria* 22 (Summer 1986), pp. 95-106. The role of the National Archives of Canada in making its records available for the Japanese-Canadian redress programme is explicitly dealt with in Nancy McMahon, "Coming Full Circle: Contemporary Uses of the Records of the Office of the Custodian of Enemy Property," paper delivered at the Annual Conference of the Association of Canadian Archivists, Victoria, BC, 1 June 1990.
  - 4 Danielle Laberge, "Information, Knowledge, and Rights: The Preservation of Archives as a Political and Social Issue," *Archivaria* 25 (Winter 1987-88), pp. 44-50.
  - 5 It is important to emphasize that most public policy research uses information **extracted** from case records, and therefore is interested more in runs of data rather than in series of records. Similarly, the research methods used by sociologists, public policy-makers, and social historians to evaluate case file information are **not** equivalent to the research and appraisal methods of archivists, although there may be a useful cross-fertilization.
  - 6 Parr, "Case Records," p. 136, for both quotations.
  - 7 This was the central contention in the Federal Bureau of Investigation case file incident; the FBI assertion that information on individuals held in field office case records was either duplicated at headquarters or incorporated in reports filed there was shown, after long study, to be untrue. The best, short summary of this important case is found in James Gregory Bradsher, "The FBI Records Appraisal," *The Midwestern Archivist* 13 (1988), pp. 51-66.
  - 8 See Duchein, *Obstacles to the Access, Use and Transfer of Information From Archives*, for an excellent summary of this problem. A good introduction to the issues involved is D.H. Flaherty, *Protecting Privacy in Surveillance Societies: The Federal Republic of Germany, Sweden, France, Canada, and the United States* (Chapel Hill, NC, and London, 1989).
  - 9 On this last issue relating to the prosecution of alleged war criminals, see Terry Cook, "Nazi Cases Not A Factor. For the Record: Archivists Honourable," *The Globe and Mail*, Toronto, 11 August 1986, p. A7; and Robert Hayward, "'Working in Thin Air': Of Archives and the Deschênes Commission," *Archivaria* 26 (Summer 1988), pp. 122-36.
  - 10 These assertions are supported in detail in the theoretical companion piece to the present article; see Terry Cook, "Mind Over Matter: Towards a New Theory of Archival Appraisal," to be published in the *festschrift* for Hugh Taylor in 1992. That piece outlines an appraisal model to determine which of the thousands of series may have archival value; the present essay offers guidelines of how to sample or select for preservation portions of those series determined thereby to have such value.

- 11 F. Gerald Ham, "Archival Choices: Managing the Historical Record in an Age of Abundance," in Nancy E. Peace, ed., *Archival Choices: Managing the Historical Record in an Age of Abundance* (Lexington, Mass., 1984), p. 133.
- 12 At the National Archives of Canada, for example, my study, "The Appraisal of Case Files: Sampling and Selection Guidelines for the Government Archives Division," was based in part on four separate studies of sampling previously commissioned by the Division over the past decade. Despite the real value of these earlier reports, the Division did not feel that, alone, they gave archivists the range of appraisal tools needed to cope with case files. The focus of the problem must be appraising case files, not sampling as one means whereby an appraisal decision may be implemented.
- 13 See, for example, Paul Lewinson, "Archival Sampling," *American Archivist* 20 (October 1957), pp. 291-312; or Felix Hull, *The Use of Sampling Techniques in the Retention of Records: A RAMP Study with Guidelines* (Paris, 1981), p. 10, and *passim*.
- 14 See F. Gerald Ham, "The Archival Edge," in Maygene F. Daniels and Timothy Walch, eds., *A Modern Archives Reader* (Washington, 1984), p. 326. Richard Berner believes that appraisal theory is so primitive that he deliberately left it out (see pp. 6-7) of his *Archival Theory and Practice in the United States: A Historical Analysis* (Seattle and London, 1983), adding that "a body of appraisal theory is perhaps the most pressing need in the archival field today."
- 15 The most important statement (from 1972 originally, and reflecting in its text and notes the debate in Europe at that time) is Hans Booms, "Society and the Formation of a Documentary Heritage: Issues in the Appraisal of Archival Sources," *Archivaria* 24 (Summer 1987), pp. 69-107, translated by Hermina Joldersma and Richard Klumpenhouwer (who provide a brief introduction).
- 16 I have tried to suggest how this integration of structure, function and participant might occur, and have advanced one possible integrationist appraisal model; see "Mind Over Matter: Towards a New Theory of Archival Appraisal," *passim*; and, for more detail, *The Archival Appraisal of Records Containing Personal Information*, Chapter 3. Both these pieces also explore in considerable detail the failings of current North American appraisal theory alluded to above.
- 17 For more details of the planned approach and its criteria, see National Archives of Canada, "Government-Wide Plan for the Disposition of Records, 1991-1996," internal report (November 1990), pp. 6-16. I wrote this report, with the aid of an advisory committee. For the various agendas of the archivist and his or her strategic alliances, see my "Mind Over Matter."
- 18 The concept and term are Trudy Peterson's, in her letter to the author, 19 March 1990.
- 19 See National Archives and Records Administration, Office of Records Administration, *Appraisal of Department of Justice Litigation Case Files: Final Report* (Washington, 1989), *passim*. This published report of under fifty pages is an excellent, concise example of appraising case records and its methodology will interest readers of this article. The Justice model followed that of the FBI case, which was in turn patterned after the Massachusetts court records project led by Michael Hindus. (See Bradsher, "FBI Records Appraisal," pp. 55-56, and note 38 below.)
- 20 For more discussion, see my *Archival Appraisal of Records Containing Personal Information*, section 2.22 and 2.23. Not all forms of all the records in these categories need be retained.
- 21 An example from the Public Record Office involving 300,000 feet (to 1954) of shipping and seamen's records, which were divided and shared with several repositories even crossing national borders, is described in Michael Cook, *Archives Administration* (London, 1977), pp. 73-74.
- 22 Jenny Dean and Wendy Southern, "The Practice of Sampling in the Disposal of Commonwealth Records," *Archives and Manuscripts* 18 (May 1990), p. 62.
- 23 This apparently popular expression among statisticians was conveyed to Rod Young, whom I thank for the reference, by Rick Ciok, Small Area Data Division, Statistics Canada.
- 24 The leading expert on archival sampling concluded his report on the subject by calling sampling "the worst of all worlds," adding "it should not be adopted unless there is no alternative solution. . ." See Hull, *Use of Sampling Techniques*, p. 55.
- 25 Statisticians are very sceptical of the use of such non-probability methods for reliable statistical analysis: see Robert Mason, *Statistical Techniques in Business and Economics* (Homewood, Ill., 1982), p. 308; Hubert M. Blalock, Jr., *Social Statistics* (New York, 1972), pp. 527-28; Lesley Kish, *Survey Sampling* (New York, 1965), pp. 18-19, 28-29; Russell Langley, *Practical Statistics for Non-Mathematical People* (New York, 1971), pp. 49-50; and David S. Moore, *Statistics: Concepts and Controversies* (San Francisco, 1979), p. 6-7. These references, and those in

- following footnotes from books by statisticians, are based on Tom Nesmith's work for the Government Archives Division. As mentioned in note 1 above, he left a core of research on sampling before departing from the National Archives of Canada, which I was then asked to expand considerably and complete.
- 26 See Jake V. Th. Knoppers, "Report on Archival Sampling Strategy and Related Issues," a contract study presented October 1983 to the National Archives of Canada, where this method was advanced and explained.
  - 27 Beyond these conceptual issues, there are practical problems as well. The alleged principal merit of the "F" sample was the possibility (highly specious, anyway, for hard-copy case files) of linking the various "F" files from all archivally valuable series across the entire government. This was undermined by the fact that a great many case file series are not organized or labelled alphabetically. The same problem occurs when information on many individual Canadians is "batched" in one single file, usually organized and/or labelled by date, location or function, and rarely alphabetically. Records disposal personnel also found pulling the "Fs" for any series other than those arranged and labelled alphabetically by surname, to be an extremely time-consuming task.
  - 28 Blalock, *Social Statistics*, p. 45.
  - 29 *Ibid.*, p. 516.
  - 30 Langley, *Practical Statistics*, p. 47.
  - 31 Kish, *Survey Sampling*, p. 25. This paragraph on sample size is especially indebted to Tom Nesmith's work.
  - 32 This contrasts with the recommendation of Felix Hull, in *Use of Sampling Techniques*, p. 16.
  - 33 The SIN terminal digit 5 method was advanced by Jake Knoppers, in his "Report on Archival Sampling Strategy and Related Issues." He analysed the geographical and mathematical properties of the social insurance number, before coming to this conclusion. While still useful as noted for some small series, this method has gradually been abandoned at the National Archives because, in addition to problems of sample size and the difficulty of retrieving it, the "linkage" possibilities across series often did not materialize (many cases were batched, or never used the SIN as a file identifier); it was of no use for records predating 1964 (the introduction of the SIN); and records managers are now increasingly reluctant to use the SIN as a file designator for private citizens because of privacy considerations.
  - 34 I received helpful advice for this section on the types of sampling from Tina Lloyd, National Archives of Canada, whose statistical knowledge far exceeds my own.
  - 35 An analogous case involved retaining permanently all (rather than a second sample) of the small number of conscientious objector cases from a large series of British appellate tribunal files relating to military service call-up. This was done after a systematic sample was taken of the entire series, which — with the exception of the conscientious objectors — was overwhelming routine, consisting of brief time extensions to tradespeople allowing them to put their affairs in order before call-up to the armed forces. See Hull, *Use of Sampling Techniques*, pp. 12-13.
  - 36 While details and examples have been added, and the terminology changed, my discussion of the advantages and disadvantages of various sampling and selection methods here and later is modelled on Trudy Huskamp Peterson, "Summary of Sampling Techniques," in her *Basic Archival Workshop Exercises* (Chicago, 1982), pp. 12-13, and is used with her permission.
  - 37 Lesley Kish notes that "determining these boundaries may prove a subjective and worrisome task." See *Survey Sampling*, p. 119.
  - 38 Archivists wishing to explore sampling more fully by looking at particular cases are referred to three good studies: National Archives and Records Service, *Appraisal of the Records of the Federal Bureau of Investigation: A Report to Hon. Harold T. Greene, U.S. District Court for the District of Columbia* (Washington, 1981); Michael Stephen Hindus, Theodore M. Hammett, and Barbara M. Hobson, *The Files of the Massachusetts Superior Court, 1859-1959: An Analysis and a Plan for Action* (Boston, 1979); and the NARA Department of Justice appraisal described in note 19 above. The most recent major study is Comité interministériel sur les archives judiciaires (Quebec), *Rapport du sous-comité sur l'échantillonnage* (Montreal, 1989), and it is recommended; an abridged English-language version has also been published: *Report of the Interministerial Committee on Court Records* (Montreal, 1991). The conclusions concerning sampling in some of these studies do not necessarily accord with those in this article.
  - 39 National Archives and Records Service, *Disposition of Federal Records* (Washington, 1981), table 4, cited in Leonard Rapport, "In the Valley of Decision: What To Do about the Multitude

of Files of Quasi Cases," *American Archivist* 48 (Spring 1985), p. 178, n. 10. Rapport raises doubts in his article about whether such criteria do not still bring too many useless records into archives.

- 40 Common sense must prevail here. Some institutions, like the police or courts, by definition only deal with "problem" cases. The point is to see whether the problem cases for a certain programme or function are centralized, which *ipso facto* segregates them from the vast majority of routine "non-problem" cases surrounding them. If so, then the vast bulk of the non-problem case files left behind may well be destroyed.
- 41 The concept was also used in the FBI appraisal case, as well as in other investigations of sampling. For an analysis of the value of the "fat file" syndrome, see NARA, *Appraisal of Department of Justice Litigation Case Files*, pp. 47-49, and *passim*.
- 42 For a fuller although somewhat dated discussion, see Hull, *Use of Sampling Techniques*, pp. 35-37.
- 43 For a description of the special characteristics of electronic records and their significance for appraisal, see Harold Naugler, *The Archival Appraisal of Machine-Readable Records: A RAMP Study With Guidelines* (Paris, 1984), pp. 37-41.