Lot sampling plans in the measure of quality of care indicators

ANTONI CORBELLA JANÉ¹ AND PERE GRIMA CINTAS²

¹Institut d'Estudis de la Salut, Departament de Sanitat i Seguretat Social and ²Escola Tècnica d'Enginyers Industrials, Departament d'Estadística i Investigació Operativa, Universitat Politècnica de Catalunya, Barcelona, Spain

Abstract

Objective. Application of the lot sampling technique to measure indicators of quality of care on screening activities directed toward women: detection of risk factor for cardiovascular diseases, screening for breast cancer and for cervical cancer.

Design. The sampling plan adopted involved, in each centre, a sample size of 15 and a maximum of four non-compliers that the random sample would tolerate. The binomial distribution was used. This 15–4 sampling plan sets the probability of concluding that an indicator is performing well for a lot or population that has an 80% compliance level at 0.84, keeping the α error or provider's risk at 0.16, while those for a population that has a 50% compliance is set at only 0.059, keeping the β error or consumer's risk at 0.059.

Setting. Ten primary care centres of the Catalan Health Institute.

Patients. Random sample of all women aged 40–65 years enrolled in each centre.

Results. For indicators of risk factors for cardiovascular disease detection, six out of 10 centres performed well for hypertension and obesity, and five out of 10 centres performed well for hypercholesterolaemia. However, no centre was acceptable either in cervical cancer or breast cancer screening.

Discussion. Rather than seeking to obtain precise estimates, this technique aims to facilitate the decision-making process regarding the quality levels of the indicators examined. Despite some limitations, the technique is a good test for detecting gross departures from stated compliance thresholds. It has several advantages for application in health services: (i) the use of small sample sizes; (ii) the fact that it can detect deficiencies in small areas; and (iii) the simplicity of applying this technique by relying on ready-made tables.

Keywords: lot sampling plans, quality improvement, quality indicators, screening activities in women

In the health care management field, measuring compliance levels of quality indicators is an issue that is of interest to both external quality oversight and internal quality activities. However, the measurement of quality indicators must strike a balance between what is desirable and what is attainable, particularly when quality improvement plans include periodic or routine follow-up as part of health care tasks. Health care quality surveillance that is based on precise estimates requires significant resource investment. For example, if in a large population one sets out to attain a 95% confidence interval (CI) in the results from measuring qualitative variables, samples of roughly 400 cases will be needed [1]. For health services institutions, this would entail a level of resource allocation that is difficult to maintain, at least on a routine basis.

To address this issue, the continuing education activities in quality improvement that have been organized for health care professionals by the Institute for Health Studies in collaboration with the Department of Statistics and Operational Research at the Polytechnic University of Catalonia, have led to quality improvement projects in which the lot sampling technique has been applied.

Background

The lot sampling technique emerged in industrial quality management [2,3] where it has been used for decades in the so-called in-coming inspection, in order to accept or reject homogeneous product lots based on previously established quality levels. Rather than seeking to obtain precise estimates, this technique aims to facilitate the decision-making process regarding the quality of examined lots.
Within health services, monitoring and evaluation through the use of lot sampling plans or the so-called lot quality acceptance sampling, is akin to conducting cross-sectional studies. Other methods that also originated in the industrial world, such as control charts used in statistical process control, would be equivalent to the ongoing surveillance of health service processes [4]. Curiously, although introduced more than 60 years ago, they have not been used in the field of health service quality management until recently.

Methodological basis of the lot sampling plans
A brief description of the lot sampling technique is given in Appendix 1; more information about it can be found elsewhere [4–6]. Basically, it entails:

(i) establishing the health unit or service that will be evaluated;
(ii) establishing and defining the quality specifications or indicators that will be reviewed;
(iii) defining what is understood as a lot: in terms of the health services, a lot could be defined as a set of patients cared for by a given service or under the care of a given professional;
(iv) establishing a sampling plan. In order to come up with a sampling plan, the following must be established:

- upper and lower thresholds. The upper threshold or Acceptable Quality Level (AQL), and lower threshold or Rejectable Quality Level (RQL);
- accepted and agreed-upon values for consumer’s risk and provider’s risk.

As in other traditional sampling methods, lot sampling plans accept some risks: a lot may be rejected as defective when, in fact, it is acceptable: provider’s risk (type I or α error); and a lot may be accepted as good when, in fact, it is defective: consumer’s risk, (type II or β error). These four values determine the sample size (n) and the maximum number of defective cases or non-compliers that the sample will tolerate (d). At its most basic, the sampling plan is determined by these two values: n and d. For example, a 10–3 sampling plan means that an n = 10 random sample has been taken. If the number of non-compliers is 0, 1, 2 or 3, the batch will be classified as acceptable and the consumer must accept the batch. However if the number of non-compliers is greater than 3, the batch will be classified as defective and will be rejected.

The lot sampling plans technique operates on the basis of probability calculations outlined in the Appendix. These calculations are time-consuming, although they can be made easier by relying on statistical software packages or tables such as binomial cumulative probabilities or those published by Valadez [6].

Graphically, constructing the so-called characteristic operating curve (Figure 1) shows the RQL, the AQL, and the provider’s and consumer’s risk values. In this figure, the percentage of compliance is on the horizontal axis and the acceptance probability for the lot is on the vertical axis. The values for the two compliance thresholds (AQL and RQL), and for the two stipulated risks (consumer’s and provider’s risk), establish two points intercepted by a single characteristic operating curve.

Material and methods
Health services evaluated
In Spain, almost 100% of the population is covered by the public health care system. The governmental Catalan Health Service (Servei Català de la Salut; SCS) finances public primary care services. Primary care centres constitute the basic structure where primary care teams develop their professional activity providing care to communities with 5000 to 25 000 inhabitants depending on the geographical area; about 2000 patients are assigned to each physician. The teams are formed by general practitioners, paediatricians, nurses, and administrative personnel as well as by a dentist and a social worker in some of the centres. They provide care for patients both at the centre and at home. Professionals on the teams are salaried employees. In some centres a new model in managing services has been initiated in which professionals are appointed as independent enterprises providing primary care services for the Catalan Health Service.

This evaluation project was carried out in 1995 in 10 primary care centres that are part of the governmental Catalan Health Institute (Institut Català de la Salut; ICS) and that participate in the quality improvement continuing education programme of the Institute of Health Studies [7,8]. Health centre professionals carried out the evaluation.

Indicators reviewed
Selection of the clinical indicators was based on the proposal from the Spanish Society for Family and Community Medicine [9] for their preventive programme...
activities. The five items selected focus on detecting risk factors for cardiovascular disease in women aged 40–65 years: hypertension, obesity, and hypercholesterolaemia, as well as cervical cancer screening and breast cancer screening. This sampling frame made it possible to apply the five selected indicators to the same, single sample. The time window observed was between 2 and 6 years, depending on which indicator was reviewed.

**Defining lots**

Lots to be sampled were defined as the records of all women aged 40–65 years who are enrolled in each centre. In three centres, only patients under the care of a portion of the roster of primary care physicians could be studied. Cases were selected through simple random sampling.

**Sampling plan adopted**

The selection of the sampling plan followed the steps described earlier:

(i) defining the lower and upper quality thresholds: these thresholds were selected to set RQL at 50%, and AQL at 80%;

(ii) defining the consumer’s and provider’s risk. These risks were defined to set consumer’s risk at about 0.05, and provider’s risk at about 0.15;

(iii) selecting the sampling plan: sample size \( n \) to be examined, and the maximum number of non-compliers \( d \) that the sample will tolerate. By relying on binomial cumulative probabilities [6], the sampling plan adopted involved a sample size \( n = 15 \), and a maximum of defects or non-compliers that the sample would tolerate \( d = 4 \).

Based on this probability distribution, the resulting sampling plan 15–4, establishes the acceptance probabilities at 0.84 for a lot or population with an 80% compliance level, whereas those for a lot that has a 50% compliance level are set only at 0.059. To summarize, quality levels were set at above 50%, allowing for a 0.16 provider’s risk in an 80% compliance level and a 0.059 consumer’s risk in a 50% compliance level.

It may be useful for the reader to be provided with examples of other sampling plans with their RQL and AQL values, and associated risks. These examples are shown in Table 1.

**Results**

Lot acceptance in risk factors for cardiovascular disease detection was six out of 10 health centres for hypertension and obesity and five out of 10 for hypercholesterolaemia. However, no lot was accepted in either cervical cancer or breast cancer screening (Table 2). Five health centres were accepted as adequately performing for all three risk factors for cardiovascular disease indicators, one centre was accepted only for two of them, and four of them were rejected for all five indicators reviewed. In summary, risk factors for cardiovascular disease detection were acceptable in 17 out of 30 reviews, whereas cervical cancer and breast cancer detection were not acceptable in any of 20 reviews.

---

**Table 1** Characteristics of different sampling plans based on probability calculations

<table>
<thead>
<tr>
<th>( n - d )</th>
<th>RQL–AQL</th>
<th>( \beta ) Error or consumer’s risk</th>
<th>( \alpha ) Error or provider’s risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>15–4</td>
<td>50–80%</td>
<td>5.9</td>
<td>16.4</td>
</tr>
<tr>
<td>16–5</td>
<td>10.5</td>
<td>8.2</td>
<td>10.6</td>
</tr>
<tr>
<td>17–5</td>
<td>7.2</td>
<td>10.6</td>
<td>5.1</td>
</tr>
<tr>
<td>18–6</td>
<td>11.9</td>
<td>5.1</td>
<td>8.4</td>
</tr>
<tr>
<td>19–6</td>
<td>8.4</td>
<td>6.8</td>
<td>5.8</td>
</tr>
<tr>
<td>20–6</td>
<td>8.4</td>
<td>6.8</td>
<td>3.9</td>
</tr>
<tr>
<td>21–6</td>
<td>6.7</td>
<td>5.6</td>
<td>4.7</td>
</tr>
<tr>
<td>22–7</td>
<td>4.7</td>
<td>7.2</td>
<td>11.9</td>
</tr>
<tr>
<td>23–7</td>
<td>7.6</td>
<td>3.6</td>
<td>5.4</td>
</tr>
<tr>
<td>24–8</td>
<td>5.4</td>
<td>4.7</td>
<td>3.8</td>
</tr>
<tr>
<td>25–8</td>
<td>4.7</td>
<td>4.7</td>
<td>3.8</td>
</tr>
<tr>
<td>26–8</td>
<td>3.8</td>
<td>5.9</td>
<td>6.1</td>
</tr>
<tr>
<td>27–9</td>
<td>6.1</td>
<td>3.0</td>
<td>4.4</td>
</tr>
<tr>
<td>28–9</td>
<td>4.4</td>
<td>3.9</td>
<td>3.9</td>
</tr>
<tr>
<td>29–9</td>
<td>3.1</td>
<td>4.9</td>
<td>3.1</td>
</tr>
<tr>
<td>30–10</td>
<td>4.9</td>
<td>2.6</td>
<td>3.1</td>
</tr>
<tr>
<td>15–3</td>
<td>60–90%</td>
<td>9.1</td>
<td>5.6</td>
</tr>
<tr>
<td>16–3</td>
<td>6.5</td>
<td>6.8</td>
<td>4.6</td>
</tr>
<tr>
<td>17–3</td>
<td>9.4</td>
<td>2.8</td>
<td>3.3</td>
</tr>
<tr>
<td>18–3</td>
<td>7.0</td>
<td>3.5</td>
<td>5.1</td>
</tr>
<tr>
<td>18–4</td>
<td>4.7</td>
<td>9.8</td>
<td>5.6</td>
</tr>
<tr>
<td>19–4</td>
<td>9.4</td>
<td>2.8</td>
<td>3.3</td>
</tr>
<tr>
<td>20–4</td>
<td>9.4</td>
<td>2.8</td>
<td>5.1</td>
</tr>
<tr>
<td>21–4</td>
<td>7.2</td>
<td>1.8</td>
<td>3.7</td>
</tr>
<tr>
<td>21–5</td>
<td>5.4</td>
<td>2.3</td>
<td>9.6</td>
</tr>
<tr>
<td>22–5</td>
<td>4.0</td>
<td>2.8</td>
<td>5.4</td>
</tr>
<tr>
<td>23–5</td>
<td>2.9</td>
<td>3.3</td>
<td>2.9</td>
</tr>
<tr>
<td>24–5</td>
<td>5.6</td>
<td>1.2</td>
<td>4.4</td>
</tr>
<tr>
<td>25–5</td>
<td>4.2</td>
<td>1.5</td>
<td>3.1</td>
</tr>
<tr>
<td>26–6</td>
<td>2.3</td>
<td>2.2</td>
<td>4.2</td>
</tr>
<tr>
<td>27–6</td>
<td>4.4</td>
<td>0.8</td>
<td>2.3</td>
</tr>
<tr>
<td>28–6</td>
<td>11.2</td>
<td>7.5</td>
<td>11.2</td>
</tr>
<tr>
<td>29–6</td>
<td>9.0</td>
<td>8.9</td>
<td>9.0</td>
</tr>
<tr>
<td>30–7</td>
<td>7.2</td>
<td>10.5</td>
<td>4.4</td>
</tr>
<tr>
<td>31–7</td>
<td>5.7</td>
<td>12.2</td>
<td>10.9</td>
</tr>
<tr>
<td>32–11</td>
<td>10.9</td>
<td>6.2</td>
<td>5.7</td>
</tr>
<tr>
<td>33–11</td>
<td>8.8</td>
<td>7.4</td>
<td>8.8</td>
</tr>
<tr>
<td>34–11</td>
<td>7.1</td>
<td>8.8</td>
<td>7.1</td>
</tr>
</tbody>
</table>

\( n \), Sample size; \( d \), maximum of non-compliers that the sample will tolerate; RQL, Rejectable Quality Level or lower threshold; AQL, Acceptable Quality Level or upper threshold.
Among women aged 40–65 years who are enrolled in the centre, the following procedure must have been performed: 1 blood pressure check within the last two years (1993–1994); 2 weight check within the last 4 years; 3 cholesterol level check within the last 6 years; 4 In women aged 40–49 years, the following must have been performed: clinical breast examination within the last year and a mammogram within the last 2 years; 5 In women aged 40–65 years, the following must have been performed: two Papanicolaou smears for 2 consecutive years, regardless of which years, and subsequently, one test within the last 5 years.

### Discussion

Two types of discussions emerge from this experience: one deals with aspects specific to the lot sampling application in this project and the other, more general one, deals with using this technique to measure quality indicators in health care services.

#### Implications from the application of lot sampling technique in this evaluation

A sampling plan 15–4 was selected for this work, so that it could discriminate between the quality levels of the two sets of indicators of screening activities examined: risk factors for cardiovascular disease detection and cancer detection. Before this study was conducted, health professionals were under the impression that cancer detection activities were carried out at lower quality levels than were risk factors for cardiovascular disease detection activities. Results from this work confirm that view.

Comparisons with other situations could not be made. Most of the scanty published reports using this technique in health services have been carried out on vaccination coverage in developing countries [10–15]. As with other evaluations of the quality of screening activities conducted with traditional measurement methods, this survey revealed suboptimal levels in the conduct of this area, in this instance, regarding screening activities aimed at women. These activities, therefore, can be considered as problem areas that should be targeted for improvement.

The experience and results obtained in this project will become truly useful if they lead professionals to reflect on and critique the causes for failure (which frequently vary from site to site). This critical analysis should, in turn, lead to the implementation of measures designed to correct structure and process — within each centre’s limitations — and to improve the quality levels examined in this review. Health centres participating in this project have committed to re-evaluate the situation to assess the extent to which improvements have been made.

**Implications from using lot sampling to measure quality of care indicators**

As with other more traditional sampling methods, it should be noted that a basic issue could emerge here that is often not given the attention it deserves: the importance of selecting cases strictly on a random basis. Any case selection that deviates from this will invalidate the technique’s usefulness. It is advisable that the selection process — simple random, systematic — be clearly stated beforehand, and that its execution be subject to a subsequent audit for compliance.

Another feature in sampling is the importance of the homogeneity of the population subjected to inspection because the decision adopted, based on the sample, will affect the entire population. For example, consider a population in which one portion has many non-compliers and another part has just a few. A decision that assumes that the number of non-compliers requires rejection of the lot would be correct for the first portion, but incorrect for the second.

We wish to highlight the following advantages, limitations, and aspects that require further comments on this technique.

**Using small sample sizes: an alternative measurement technique to those traditionally used**

The capability of measuring quality levels in a different way than that offered by methods traditionally used in evaluating the health services is noteworthy. Traditional methods set out to obtain precise measurements, which require large sample sizes that, in turn, lead to a review of the clinical records or other documentation of many cases.

**Identifying deficiencies in small areas**

Another noteworthy feature of the lot sampling technique is its ability to point out possible flaws in small areas, such as a single health centre or a hospital department. This holds promise for quality management at the local level, which frequently shows problems that vary according to such factors as each site’s structure and organization of processes. If one applies traditional sampling methods, this is made difficult, given that the sample size that must be used to arrive at the
Lot sampling plans in evaluation of health services

conventional degree of 95% CI is very large. This implies that an arduous review of vast amounts of information must be carried out [16], which places traditional methods beyond the possibility of any routine use.

Lot sampling does not produce precise estimates

The use of small sample sizes entails a disadvantage that offsets an advantage noted in the previous section: this technique cannot obtain precise values from its measurement.

Therefore, the lot sampling technique might not detect small changes in the ongoing surveillance or monitoring of accumulated indicators (on the same populations, and over time) in which large variations are not expected.

However, in this case its usefulness depends on the subject, each indicator, evaluated. For example, if the level of compliance found in the first measure is quite low then, choosing an appropriate RQL–AQL interval in the next measure, the technique will more likely detect changes. On the other hand, if the level of the first measure is quite high, then the objective of the lot sampling technique should shift to focus on the maintenance of these levels.

Finally, to detect small changes, more traditional sampling methods should be used. This is the price that using small sample sizes exacts. If precise measures are required, either lot sampling should be used as a screening technique to be supplemented subsequently with other more traditional sampling methods, or more traditional methods should be used at the outset.

The technique’s validity

The risk values of the sampling plan sets the probabilities of misclassifying a lot as acceptable when, in fact, it is poor (consumer’s risk) and as poor when, in fact, it is acceptable (provider’s risk). For example, in our survey, an indicator with an 80% compliance level in the lot or population, has an 83.6% probability of being correctly classified as acceptable; and if an indicator only achieves a 40% compliance level, it has a 94.1% probability of being classified as unacceptable. But it should be noted that the provider’s and consumer’s risk only adopt exactly these values in the unlikely event that the compliance level in the populations coincide with the AQL or RQL values. When the compliance levels in the populations are different from the RQL or AQL, which is perfectly plausible, then the risks will adopt other values as shown graphically in the operating characteristic curve. This curve plots the probabilities of accepting and rejecting the lots against the compliance rate (Figure 1). For example, in our survey if the compliance level in the population was 70%, it is about as likely to be classified acceptable as it is to be classified unacceptable. Compliance levels intermediate between the RQL and AQL will be necessarily associated with higher risks [17].

Therefore, for a health care organization to make use of the lot sampling technique, it should be emphasized that it is important to have a rough idea about the compliance levels of the indicators, before it sets out to propose the RQL–AQL levels for a specific sampling plan.

Lot sampling as an unfamiliar technique in health services

This kind of statistical technique is unfamiliar to health care professionals. From a management point of view, this problem should be handled not as a limitation of the lot sampling technique itself, but as an obstacle to its implementation and acceptance in health services. It is not easy for some to understand these relatively new approaches for measuring quality indicators or other kinds of items, especially, the first time that one hears or reads something about them. In our experience, health professionals initially express doubts about these statistical techniques of quality control. At the beginning, they are not comfortable without the traditional measures based on precise estimates. Due to their experience with traditional approaches to evaluation in health services, they tend to look for precise estimations and CIs.

For all of these reasons, to introduce and extend utilization of these techniques in health services, it will be necessary to explain the principles and provide practical examples.

Lot sampling plans have certain limitations and weaknesses, which should be known before applying this technique. It is not possible to detect small changes in the ongoing surveillance or monitoring of accumulated indicators over time, in which large variations are not expected. Despite these limitations, it appears to be a useful alternative for measuring quality of health care. It is a good test for detecting gross departures from stated compliance thresholds. In summary, the advantages for application in health care services are: (i) the use of small sample sizes; (ii) the fact that it can detect deficiencies in small areas; and (iii) the simplicity of applying this technique by relying on ready-made tables [6].

Acknowledgements

We thank the following coordinators for this project within the primary care services of the Catalan Health Institute: Mercé Bonet, Health Center (HC) La Jonquera; Ignasi Casado (HC Balsa-Lleida); Fernando Dolz (HC Falset); Olga Fernández (HC Montornès-Montmeló); Pilar García (HC Raval-Nord-Barcelona); Concepción Morera (Primary Care Subdivision-Girona); Aurora Palos (HC Borges Blanques); Montserrat Mas and Xavier Pérez-Bermejo (HC Girona 4); Josep Sabaté (HC Borges del Camp); Pasqual Solanas (HC Girona 3);Montserrat Vidal (HC Sils); and to Ronald Epstein (University of Rochester, New York) for the review of an earlier draft.

References


Appendix

Examples of probability calculations in lot sampling plans

The lot sampling technique operates on the basis of probability calculations— for example, the probability of finding \( d \) or fewer defects or non-compliers cases in a random sample of size \( n \) taken from a lot or population that shows certain compliance threshold or percentage. Using formulas of the hypergeometric, Poisson, or binomial distributions can derive this probability. For example, using binomials, the probability \( P(x) \) of finding \( x \) individuals with an event in a sample of size \( n \) is given by the expression:

\[
P(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}
\]

where: \( P(x) = \) probability of finding \( x \) compliers in a sample of \( n \) cases; \( n = \) sample size; \( x = \) number of compliers in the sample; \( p = \) expected proportion of compliers in the population; \( q = 1 - p \); and by definition \( 0! = 1 \).

For example, assume that the RQL and AQL compliance thresholds selected are 50% and 80%; and the accepted risks are about 0.05 for consumer’s risk and 0.15 for provider’s risk. Then, relying on ready-made tables of the binomial cumulative probabilities [6], an appropriate sampling plan is selected. This sampling plan was 15–4 (\( n = 15, d = 4 \)). Thus, the probability of finding four non-compliers in a sample of 15 cases, taken from a population with 50% compliance is:

\[
P(d = 4) = \frac{15!}{11!(15-11)!} 0.50^4 (0.50)^{11} = 0.041
\]

and the probability of finding no non-compliers in a sample of 15 cases, taken from a population with 80% compliance is:

\[
P(d = 0) = \frac{15!}{0!} 0.80^0 (0.20)^{15} = 0.035
\]

Accepted for publication 2 November 1998