Title
Optimal design of multiple-objective Lot Quality Assurance Sampling (LQAS) plans.

Journal
Biometrics, 75(2)

ISSN
0006-341X

Authors
Duarte, Belmiro PM
Wong, Weng Kee

Publication Date
2019-06-01

DOI
10.1111/biom.13008

Peer reviewed

Biometrics 64, 1–18 DOI: 10.1111/j.1541-0420.2005.00454.x
December 2017

Design of LQAS plans to meet diagnostic accuracy metrics

Belmiro P.M. Duarte[1,2]

[1]Centro de Investigação em Processos Químicos e Produtos da Floresta,
Department of Chemical Engineering, University of Coimbra,
Pólo II, R. Sílvio Lima, 3030-790 Coimbra, Portugal

[2]Department of Chemical and Biological Engineering,
ISEC, Polytechnic Institute of Coimbra,
R. Pedro Nunes, 3030-199 Coimbra, Portugal.

email: bduarte@isec.pt

and

Weng Kee Wong

Department of Biostatistics,
Fielding School of Public Health, UCLA,
10833 Le Conte Ave., Los Angeles, California 90095-1772, U.S.A.

email: wkwong@ucla.edu

Summary: We consider the problem of optimally designing Lot Quality Assurance Sampling (LQAS) plans for health monitoring purposes. A Mixed Integer Nonlinear Programming (MINLP) formulation is proposed to address the problem when pre-defined levels of accuracy of diagnostic metrics used to assess the programmes are imposed. The formulation is used for finding LQAS plans for different combinations of diagnostic metrics which are compared to classic plans based on purely statistical backgrounds.

Key words: Lot Quality Assurance Sampling; Optimal plans; Diagnostic metrics; Mixed Integer Nonlinear Programming.

This paper has been submitted for consideration for publication in *Biometrics*

1. Introduction

A Lot Quality Assurance Sampling (LQAS) is a classification procedure for decision making about the acceptance of a given lot. The strategy is grounded in the statistical tools developed around 1920-30 to help control the quality from a production line. LQAS was developed in the 1950's for the industry to check product quality and quickly found wide-spread applications in health care surveys. Robertson and Valadez (2006) provides a review of LQAS and illustrates how it can help decision makers classify a given population characteristic as acceptable or not. Lemeshow and Taber (1991) provides statistical details and also compares merits of having a single or double-sampling plan.

LQAS is commonly used in public health research to determine if a health policy or a community program is working for the intended purpose. Some specific applications of LQAS are monitoring immunization programmes to ascertain their cost effectiveness (Sandiford, 1993), monitoring elimination leprosy in a region (Gupte et al., 2004), examining effectiveness of community intervention programs of captia and management systems on maternal and child health behavior change (Valadez et al., 2005), assessing the prevalence of acute malnutrition (Deitchler et al., 2007; Olives et al., 2009; Olives and Pagano, 2010) and monitoring malaria outcome indicators (Biedron et al., 2010). Vanamail et al. (2006) discussed operation feasibility and implementation of LQAS as a tool for routine monitoring in the context of filariasis control programs. LQAS invariably includes design questions; given user-specified decision parameters for the problem, what are the optimal sample size and the decision rules to implement for, say, monitoring the effectiveness of a disease eradication community based program or a state-sponsored vaccination program? Interest in LQAS continues to date. For example, Olives et al. (2012) applied ideas to incorporate outcomes that have a few categories, not just two, with application to Schistosomiasis control.

The design of LQAS plans for health monitoring is similar to Acceptance Sampling plans by

variables or attributes for quality control purposes. Both can be formulated as optimization problems (Duarte and Saraiva, 2008, 2013), which frequently require the minimization of certain quantities, such as the sample size or the Average Sampling Number (ASN) subject to the constraints at the controlled points of the *operating characteristic* (OC) curve.

Very often the health monitoring programmes are intended to reach a previously defined level of a statistical measure quantifying the performance of the binary classification test. Among the vast number of diagnosis performance metrics, the most used are: (i.) *Specificity*; (ii.) *Sensitivity*; (iii.) *Negative Predicted Value* (NPV); and (iv.) *Positive Predictive Value* (PPV). For a review of the diagnostic performance metrics the reader is referred to Altman and Bland (1994); Fletcher et al. (2012) and in §2.2 we review the definitions commonly accepted in biostatistics community.

Despite of the advantages of the LQAS plans based on controlling two points lying to the OC curve which are easier to find and are tabulated, when a diagnosis accuracy metric is previously established it would be beneficial implementing LQAS plans specifically for that purpose. While the first are based on purely statistical knowledge, the second accounts for the specific goals of the monitoring health programmes. From our knowledge this topic has never been investigated, and in this paper we propose the first systematic approach for designing LQAS plans to meet pre-defined levels of diagnosis performance metrics.

The paper includes four additional sections. In Section 2, the mathematical background that supports our approach is presented. Section 3 introduces the MINLP formulation for designing LQAS plans to satisfy the OC-curve constraints and diagnosis performance criteria. In section 4 we present and compare results, and in Section 5 we conclude.

2. Mathematical background

In this section, we provide the background material required for the formulation and numerical solution of LQAS plan design problem. In section 2.1 we introduce LQAS plans, in

§2.2 we introduce the diagnosis performance metrics, in §2.3 we review the use of Gaussian Quadrature Formulas (GQF), and in section 2.4, we briefly review the fundamentals of MINLP.

2.1 *LQAS plans*

This section presents the fundamentals of Acceptance Sampling. Following convention, we use the binomial distribution to model the probability of individuals having the characteristic ξ in the population where ξ can model outcomes as "having a disease"/"having not a disease" or "having been vaccinated"/"having not been vaccinated" among others depending on the program monitoring purpose. In the former scenario "having not a disease" is represented with $\xi = 1$ and the opposite outcome by $\xi = 0$. In the latest context, $\xi = 1$ is to represent the outcome "having been vaccinated" and $\xi = 0$ the opposite result. Practically, the result of the test on each individual is either conforming or nonconforming, with the former corresponding the individual having the sought characteristic, i.e. a success with $\xi = 1$. Throughout, we make two assumptions: (i) the probability of selecting nonconforming/conforming individuals is independent of the sampling method; and (ii) the potentially infinite number of individuals of the population is not impacted by the sample.

We represent a sampling plan by $\mathcal{S}(n, r)$, where n is the sample size and r is the acceptance limit used to declare a lot/population acceptable (or not) based on a binary outcome ξ taking values 0 or 1. A primary goal is to assure that the lot or population is acceptable when the proportion of outcomes ξ meets a given conformity proportion. The procedure randomly samples n individuals and the population/lot is accepted if the number of individuals in the sample tested positive ($\xi = 1$) is greater or equal to r. Otherwise, the population/lot is declared unacceptable. Since the decision is based on testing a particular hypothesis, the inference is subject to statistical type I and II error rates. To measure the classification

errors we use the *probability of acceptance* of lots/populations with a given proportion p of successes, denoted by $P_a(p)$.

Very often the LQAS plans are designed to reach pre-specified levels of risk of misclassification of populations/lots at required lower and upper proportions of success, p_L and p_U, respectively. The constraints imposed to the plan are grounded on statistical knowledge, and require that the probability of incorrectly classify a lot of upper quality level p_U as unacceptable (type I error) should be lower than α, and the probability of incorrectly classify a lot of lower quality p_L as acceptable (type II error) should be lower than β, i.e.

$$P(x \geqslant r|p_L) \leqslant \beta \tag{1a}$$

$$\text{and} \quad P(x \geqslant r|p_U) \geqslant 1-\alpha. \tag{1b}$$

Here x is the number of individuals of the population/lot with the sought characteristic (i.e. $\xi = 1$) and $P(x \geqslant r|p_L)$ is the probability of considering the population/lot acceptable for a given proportion level of successes p_L.

The probability of obtaining a specified number x of individuals with the sought characteristic in a sample of n from a population with a proportion p having the characteristic $\xi = 1$ is modeled by the binomial distribution

$$\mathbb{P}(x) = \binom{n}{x} p^x \ (1-p)^{n-x}, \quad x = 0, \cdots, n. \tag{2}$$

If the number of conforming individuals tested (with $\xi = 1$) in the sample is less than r, we reject the population/lot and conclude that the programme was not succeed, otherwise we accept the population and infer that the programme is well succeed. In public health research, we may want to ascertain if the people are adequately vaccinated in a region, and so x in such a study is the number of people vaccinated in the sample.

The OC curve represents the probability of acceptance of populations/lots with a given proportion p of individuals with $\xi = 1$. For acceptance/rejection purposes of populations

tested for a characteristic following the binomial distribution (2), the OC curve is

$$F(p|n,r) = \sum_{x=r}^{n} \mathbb{P}(x) = \sum_{x=r}^{n} \binom{n}{x} p^x \ (1-p)^{n-x}, \tag{3}$$

which is computed using the regularized Beta function, here denoted by I, see Press et al. (1996, Chap. 6) for details:

$$F(p|n,r) = \sum_{x=r}^{n} \binom{n}{x} p^x \ (1-p)^{n-x} = I(p, n-r, r-1). \tag{4}$$

One of the early strategies used to design LQAS plans is based on the OC curve. In practice, the sample size and the acceptance constant are determined such that the conditions $F(p_L|n,r) \leqslant \beta$ and $F(p_U|n,r) \geqslant 1-\alpha$ are both validated for pre-defined levels p_L and p_U and risks α and β. In spite of the decision *grey region* $p_L \leqslant p \leqslant p_U$ being the interval where the consequences of the misclassification error have lower impact, there are recognized risks that are not accounted for in the LQAS plans. In particular, LQAS plans may have good sensitivity but not good specificity (Sandiford, 1993).

2.2 *Diagnosis performance metrics*

Diagnostic tests where the result is binary are conventionally summarized in a two-by-two table (Fletcher et al., 2012). The corresponding diagnostic accuracy tests is commonly measured by some metrics; among them are (i) Sensitivity; (ii) Specificity; (iii) PPV; and (iv) NPV. Formal definitions of the diagnosis performance metrics for binary outcome tests summarized in two-by-two contingency tables are, respectively (Griner et al., 1981):

$$\text{Spec.} = \frac{TP}{TP+FN}, \tag{5a}$$

$$\text{Sens.} = \frac{TN}{TN+FP}, \tag{5b}$$

$$\text{NPV} = \frac{TN}{TN+FN} \tag{5c}$$

$$\text{PPV} = \frac{TP}{TP+FP} \tag{5d}$$

Here, TP is the number of individuals with $\xi = 1$ that test positive (desirable outcome), FN is the number of individuals with $\xi = 1$ that test negative (undesirable outcome), FP

is the number of individuals with $\xi = 0$ that test positive (undesirable outcome), and TN is the number of individuals with $\xi = 0$ that test negative (desirable outcome). Specificity is the proportion of individuals with $\xi = 1$ correctly identified, Sensitivity is the proportion of individuals with $\xi = 0$ correctly identified, NPV the proportion of individuals that test negative correctly identified, and PPV is the proportion of individuals that test positive correctly identified. In what follows, we use the formal definitions (5) adapted to three-by-two Table 1. The elements of the three-by-two Table 1 are the number of individuals (or its proportion) that test negative or positive for different proportion of success where the grey zone is an additional scenario. Table 2 combined with equations (6) set the metrics for lots/populations with low proportion of individuals with $\xi = 1$ ($p \leqslant p_L$) and large proportion ($p \geqslant p_U$). In what is to follow, we call the proportion of individuals with $\xi = 1$ as the *anticipated prevalence*, as Lemeshow and Taber (1991). Typically, we have low anticipated prevalence rate (LAPR) populations and high anticipated prevalence rate (HAPR) populations if $p \leqslant p_L$ and $p \geqslant p_U$, respectively.

[Table 1 about here.]

[Table 2 about here.]

$$a = \int_0^{p_L} \sum_{x=r}^{n} \binom{n}{x} p^x (1-p)^{n-x} \, \mathrm{d}p \qquad b = \int_{p_L}^{p_U} \sum_{x=r}^{n} \binom{n}{x} p^x (1-p)^{n-x} \, \mathrm{d}p \tag{6a}$$

$$c = \int_{p_U}^{1} \sum_{x=r}^{n} \binom{n}{x} p^x (1-p)^{n-x} \, \mathrm{d}p \qquad d = \int_0^{p_L} \sum_{x=0}^{r-1} \binom{n}{x} p^x (1-p)^{n-x} \, \mathrm{d}p \tag{6b}$$

$$e = \int_{p_L}^{p_U} \sum_{x=0}^{r-1} \binom{n}{x} p^x (1-p)^{n-x} \, \mathrm{d}p \qquad f = \int_{p_U}^{1} \sum_{x=0}^{r-1} \binom{n}{x} p^x (1-p)^{n-x} \, \mathrm{d}p \tag{6c}$$

2.3 *Gaussian Quadrature Formulas*

Gaussian Quadrature Formulas are a class of methods that use appropriate weights and nodes to numerically integrate a complex function $f(t)$ to a high degree of accuracy. For a

one dimension integral over an arbitrary compact interval $[a, b]$, the formula is:

$$\int_a^b w(t)\ f(t)\mathrm{d}t \doteq \sum_{j=1}^{M} w_j\ f(t_j)$$

where $w(t)$ is a weighting function, and M is the number of points, also designated as nodes, used in the integration. The accuracy of the approximation of the integral as a sum depends on the selected weight w_j at the nodes t_j. A major advantage of GQF is that with judicious choices of the nodes and weights, it needs only M points to exactly integrate polynomials of degree 2 $M-1$ or less. This means that only M evaluations of the function $f(t)$ are required (Gerald and Wheatley, 1994). For $w(t) = 1$, $a = -1$ and $b = 1$ the nodes correspond to the zeros of the M^{th} order Legendre polynomials; see, for example, Atkinson (1989). For $w(t) = 1$ and an arbitrary compact interval on the real line, the weights and nodes are determined from recursive algorithms such as those presented in Davis and Rabinowitz (1984). The numerical approximations of the integrals (6) obtained using GQF are:

$$a = \frac{p_L}{2} \sum_{j=1}^{M} \left[1 - I\left(t_j \frac{p_L}{2} + \frac{p_L}{2}, n-r, r-1\right)\right] w_j \tag{7a}$$

$$b = \frac{p_U - p_L}{2} \sum_{j=1}^{M} \left[1 - I\left(t_j \frac{p_U - p_L}{2} + \frac{p_U + p_L}{2}, n-r, r-1\right)\right] w_j \tag{7b}$$

$$c = \frac{1 - p_U}{2} \sum_{j=1}^{M} \left[1 - I\left(t_j \frac{1 - p_U}{2} + \frac{1 + p_U}{2}, n-r, r-1\right)\right] w_j \tag{7c}$$

$$d = \frac{p_L}{2} \sum_{j=1}^{M} I\left(t_j \frac{p_L}{2} + \frac{p_L}{2}, n-r, r-1\right) w_j \tag{7d}$$

$$e = \frac{p_U - p_L}{2} \sum_{j=1}^{M} I\left(t_j \frac{p_U - p_L}{2} + \frac{p_U + p_L}{2}, n-r, r-1\right) w_j \tag{7e}$$

$$f = \frac{1 - p_U}{2} \sum_{j=1}^{M} I\left(t_j \frac{1 - p_U}{2} + \frac{1 + p_U}{2}, n-r, r-1\right) w_j \tag{7f}$$

where t_j are the zeros of the Legendre polynomials in $[-1, 1]$ and w_j are the weights. In all calculations presented in subsequent sections we consider $M = 20$.

2.4 *Mixed Integer Nonlinear Programming*

Mixed Integer Nonlinear Programming refers to a class of optimization problems including continuous and discrete variables and nonlinear functions in the objective function and/or the constraints. Mixed Integer Nonlinear Programs (MINLPs) arise in a wide range of applications, including chemical engineering, finance, and management. The general form of a MINLP is

$$\min_{\boldsymbol{x},\boldsymbol{y}} \; f(\boldsymbol{x},\boldsymbol{y}) \tag{8a}$$

$$\text{s.t. } h_i(\boldsymbol{x},\boldsymbol{y}) = 0, \quad \forall i \in \mathcal{E} \tag{8b}$$

$$g_i(\boldsymbol{x},\boldsymbol{y}) \leqslant 0, \quad \forall i \in \mathcal{I} \tag{8c}$$

$$\boldsymbol{x} \in \mathbf{X}, \; \boldsymbol{y} \in \mathbf{Y} \tag{8d}$$

where each function $h_i(\boldsymbol{x},\boldsymbol{y})$ and $g_i(\boldsymbol{x},\boldsymbol{y})$ is a mapping from $\mathbb{R}^n$ to $\mathbb{R}$, $\mathcal{E}$ is the set of equality constraints, $\mathcal{I}$ the set of inequalities, $\mathbf{X} \in \mathbb{R}^n$ is a continuous compact domain, $\mathbf{Y}$ is discrete domain containing integer values, $\boldsymbol{x}$ is the set of continuous variables and $\boldsymbol{y}$ the set of integer variables.

The most commonly used algorithms used to solve MINLPs are the outer approximation (Duran and Grossmann, 1986), the branch and bound (Fletcher and Leyffer, 1998) and the extended cutting plane (Westerlund and Pettersson, 1995). For the fundamentals of MINLP and the algorithms the reader is referred to Floudas (2002). All the problems addressed in the paper are solved with a branch and bound algorithm using the solver `SBB` (GAMS Development Corporation, 2013b) available within the general modeling system `GAMS 24.2.1` (GAMS Development Corporation, 2013a). `SBB` combines the standard branch and bound method known from Mixed Integer Linear Programming and a standard NLP solver supported by `GAMS 24.2.1`. Here `CONOPT` is used for solving the relaxed nonlinear programs (Drud, 1985) and `CPLEX` is used for solving local integer linear programs (GAMS

Development Corporation, 2013b). The relative tolerance used in all problems is 10^{-5}. In our design context the variables a, b, c, d, e and f are continuous; and n and r are integer. All computation in this paper were carried using on an Intel Core i7 machine (Intel Corporation, Santa Clara, CA) running 64 bits Windows 10 operating system with 2.80 GHz.

3. Optimal LQAS formulations

In this section we introduce the MINLP formulations for designing LQAS plans. In Section 3.1 we address the problem of finding a LQAS plan that assures that the conditions (1) at the controlled points of the OC curve are satisfied, and in §3.2 we consider the problem of designing plans for a combination of diagnosis performance criteria where lower bound thresholds are assumed. The former problem will be designated as the *OC curve-constrained design problem* and the later as the *performance criteria-constrained design problem*. In both cases the objective is the minimization of the sample size which has an economic impact. Typically, the algorithms used to design *OC curve-constrained* LQAS plans stand on enumerative procedures where n and r are successively iterated until the constraints at the controlled points of the OC curve are both satisfied (Lemeshow and Taber, 1991).

3.1 *Formulation for OC curve-constrained LQAS plans*

We consider that the risks α and β and the target proportions required for $LAPR$ and $HAPR$ populations are imposed. The resulting MINLP minimizes the sample size providing that the conditions (1) are satisfied. Equation (4) is used to represent the OC curve. The optimization

problem is as follows:

$$\min_{n,r} n \tag{9a}$$

$$\text{s.t. } I(p_U, n-r, r-1) \geqslant 1-\alpha \tag{9b}$$

$$I(p_L, n-r, r-1) \leqslant \beta \tag{9c}$$

$$n \geqslant 2 \tag{9d}$$

$$r \geqslant 1 \tag{9e}$$

$$n,\ r \in \mathbf{N} \tag{9f}$$

where equations (9b-9c) are the constraints at the OC points, (9d) and (9e) are lower bounds for n and r imposed by rational reasons.

3.2 *Formulation for performance-constrained LQAS plans*

Here, we consider that the targets for some of the diagnosis performance criteria listed in §2.2 are known. Let us designate the target for Sensitivity for populations with a $LAPR$ of $\xi = 1$ as σ_L, the target to apply in populations with $HAPR$ of individuals with $\xi = 1$ as σ_U; the target for Specificity for $LAPR$ environments as θ_L, the target for $HAPR$ scenarios as θ_U; the target for PPV for $LAPR$ environments as π_L and the target for $HAPR$ environments as π_U. Finally, we designate the target for NPV for $LAPR$ environments as ϱ_L and the target for $HAPR$ environments as ϱ_U. The reformulation of the criteria in Table 2 to avoid fractionary

terms produces the performance constraints later included in the design problem:

$$\text{Sensitivity for } LAPR \qquad a \geqslant \sigma_L(a+d) \tag{10a}$$

$$\text{Sensitivity for } HAPR \qquad f \geqslant \sigma_U(c+f) \tag{10b}$$

$$\text{Specificity for } LAPR \qquad e+f \geqslant \theta_L(b+c+e+f) \tag{10c}$$

$$\text{Specificity for } HAPR \qquad a+b \geqslant \theta_U(a+b+c+d) \tag{10d}$$

$$\text{PPV for } LAPR \qquad a \geqslant \pi_L(a+b+c) \tag{10e}$$

$$\text{PPV for } HAPR \qquad f \geqslant \pi_U(d+e+f) \tag{10f}$$

$$\text{NPV for } LAPR \qquad e+f \geqslant \varrho_L(d+e+f) \tag{10g}$$

$$\text{NPV for } HAPR \qquad a+b \geqslant \varrho_U(a+b+c) \tag{10h}$$

Following, the relations (7) together with a combination of constraints (10) chosen to construct LQAS plans to meet a given combination of diagnosis accuracy metrics are used to formulate the *performance criteria-constrained design problem*:

$$\min_{n,r} n \tag{11a}$$

$$\text{s.t. Equations (7)} \tag{11b}$$

$$\text{Combination of equations (10)} \tag{11c}$$

$$n \geqslant 2 \tag{11d}$$

$$r \geqslant 1 \tag{11e}$$

$$n,\ r \in \mathbf{N} \tag{11f}$$

$$a,\ b,\ c,\ d,\ e,\ f \in [0,1] \tag{11g}$$

Let us demonstrate our approach with an example where we want to find the LQAS plan to reach a given level of Sensitivity for $LAPR$ populations, σ_L, and for $HAPR$ lots/populations, σ_U. The problem (11) includes the relations (7) and the constraint (11c) aggregates the constraints for Specificity (10a-10b). The problems for other criteria or criteria combination

are constructed similarly. Notice that in some cases the criteria might be antagonistic which may produce infeasible solutions which are also detected by the proposed formulation. This topic will be further analyzed in §4.

4. Results

In this section we present the results for the *OC curve-constrained design problem* and compare them with those obtained for the *performance criteria-constrained design problem* for different combinations of performance criteria.

To test the formulation (9) for optimally designing LQAS plans we consider a scenario where $\alpha = \beta = 0.10$ and p_L and p_U are varied in a region commonly used in practical studies. In all scenarios the difference between p_U and p_L is kept constant but different values are considered. Table 3 presents the results and we observe that as the proportions for *LAPR* and *HAPR* populations increase, larger is the ratio r/n, and more discriminant is the plan. The results obtained are in good agreement with those of Lemeshow and Taber (1991). All examples presented in following tables require less than 1.0 s of CPU time which proves the numerical efficiency of the algorithm.

Figure 1 presents the OC curves for optimal plans obtained for scenarios S1 and S5. They illustrate the constraints satisfaction for both setups, and the larger discriminant power of the plan obtained for S5. We observe that both OC curves pass below the point (p_L, β) and above the point $(p_U, 1-\alpha)$ as required by the formulation. For the error rates considered ($\alpha = \beta = 0.1$), both plans are constrained at the point (p_L, β).

[Table 3 about here.]

[Figure 1 about here.]

Table 4 presents the LQAS for several combinations of diagnosis performance criteria. The same trend observed for OC curve constrained plans applies here; as the proportions assumed

for $LAPR$ and $HAPR$ increase, the discrimination of the plans increases as well as the ratio r/n.

To analyze the optimality of the solutions of the *performance criteria-constrained design problem* let us consider the LQAS plan obtained for scenario S5 when both the criteria sensitivity for $LAPR$ and for $HAPR$ are to be met (line 5 of Table 4). The OC curve of the plan $\mathcal{S}(6,4)$ is presented in Figure 2, which also depicts the areas $a,\ b,\ c,\ d,\ e$ and f. Table 5 lists the values of the areas and the Specificity for both groups, and we observe that the constraints ($a/(a+d) \geqslant \sigma_L$ and $f/(c+f) \geqslant \sigma_U$) are satisfied.

[Table 4 about here.]

[Figure 2 about here.]

[Table 5 about here.]

Our results demonstrate that for some specific anticipated prevalence proportions (p_L and p_U), particular combinations of diagnosis performance metrics and lower bounds, the optimization problem is infeasible and there is not an LQAS plan satisfying all constraints simultaneously. When the solution of (11) can not be obtained because of the antagonistic characteristics of the constraints, the feasibility of a relaxed problem including not the integral terms is checked with another mathematical program. That is, we convert the inequalities (10) into equivalent equalities and solve a square system of 6 algebraic equations with respect to $a,\ b,\ c,\ d,\ e$ and f using the solver `CNS` (GAMS Development Corporation, 2013b) included in `GAMS 24.2.1`. `CNS` is for constrained nonlinear systems and uses the nonlinear programming solver `CONOPT`.

Besides the relations between the parameters required by the constraints on the diagnosis metrics, others derived from geometrical assumptions using Figure 2, e.g. $a+d=p_L$, $b+e=p_U-p_L$, $c+f=1-p_U$ and $a+d+b+e+c+f=1$, can be derived. The possible results of the procedure are: (i) a combination of $a,\ b,\ c,\ d,\ e$ and f exists and the original design

problem is feasible; and (ii) there is not a combination of $a,\ b,\ c,\ d,\ e$ and f and the design is infeasible. In this case, lower values of threshold bounds must be used.

To demonstrate the use of this check tool let us consider the LQAS plan obtained for the setup where the design criteria are the Specificity for $LAPR$ and Specificity for $HAPR$ (second line in Table 4) and the threshold level is 0.95 for both criteria, i.e. $\theta_U = \theta_L = 0.95$, $p_L = 0.6$ and $p_U = 0.9$. The mathematical programming problem solved to check the feasibility of the design problem is

$$\text{find } a,\ b,\ c,\ d,\ e,\ f \tag{12a}$$

$$\text{s.t } a + b = \theta_L(a + b + d + e) \tag{12b}$$

$$e + f = \theta_U(b + c + e + f) \tag{12c}$$

$$a + d = p_L \tag{12d}$$

$$b + e = p_U - p_L \tag{12e}$$

$$c + f = 1 - p_U \tag{12f}$$

$$a + b + c + d + e + f = 1 \tag{12g}$$

$$a,\ b,\ c,\ d,\ e,\ f \in [0, 1]. \tag{12h}$$

The model (12) is infeasible, and consequently, an LQAS plan satisfying the constraints cannot be found. In practice, the maximum value of θ_U and θ_L that make the problem feasible for $\theta_L = \theta_U$ is 0.7692. If we allow $\theta_U \neq \theta_L$, then for $\theta_U = 0.95$, the maximum value of θ_L that produces a feasible plan is 0.6888. These findings are dependent of the values assumed for p_L and p_U.

5. Conclusions

We propose MINLP formulations to handle the problem of designing LQAS plans for implementing in health monitoring programmes. The design problem consists of minimizing the sample size such that a set of constraints are satisfied. First, we consider the design problem

using the classic framework where the constraints result from the points controlled of the OC curve. Next, we propose a formulation where the constraints result from a combination of diagnosis performance criteria the plan are sought to guarantee. As far as we know, our formulation is the first that can be used to design LQAS to meet pre-defined levels of diagnosis performance criteria. The later formulation requires numerically calculating integrals and we use 20-point based GQF for such a purpose. We test our proposed formulations for a large range of setups and diagnosis accuracy criteria combinations, and compare the results. Finally, we analyze cases where the combination of diagnosis performance criteria can not be satisfied simultaneously, and a feasible plan can not be found. In this case, we propose a linear programming formulation to check the feasibility of the LQAS design problem for that combination of anticipated prevalence proportions and threshold bounds.

References

Altman, D. G. and Bland, J. M. (1994). Diagnostic tests 2: Predictive values. *BMJ* **309,** 102.

Atkinson, K. E. (1989). *An Introduction to Numerical Analysis.* John Willey & Sons, New York, 2nd. edition.

Biedron, C., Pagano, M., Hedt, B., Kilian, A., Ratcliffe, A., Mabunda, S., and Valadez, J. (2010). An assessment of lot quality assurance sampling to evaluate malaria outcome indicators: extending malaria indicator surveys. *International Journal of Epidemiology* **39,** 72–79.

Davis, P. J. and Rabinowitz, P. (1984). *Methods of Numerical Integration.* Academic Press, New York, 2nd. edition.

Deitchler, M., Valadez, J. J., Egge, K., Fernandez, S., and Hennigan, M. (2007). A field test of three LQAS designs to assess the prevalence of acute malnutrition. *International Journal of Epidemiology* **36,** 858–864.

Drud, A. (1985). `CONOPT`: A GRG code for large sparse dynamic nonlinear optimization problems. *Mathematical Programming* **31,** 153–191.

Duarte, B. P. M. and Saraiva, P. M. (2008). An optimization based approach for designing attribute acceptance sampling plans. *International Journal of Quality and Reliability Management* **25,** 824–841.

Duarte, B. P. M. and Saraiva, P. M. (2013). Optimal design of acceptance sampling plans by variables for nonconforming proportions when the standard deviation is unknown. *Communications in Statistics - Simulation and Computation* **42,** 1318–1342.

Duran, M. A. and Grossmann, I. E. (1986). An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical Programming* **36,** 307–339.

Fletcher, R. and Leyffer, S. (1998). Numerical experience with lower bounds for MIQP branch-and-bound. *SIAM Journal on Optimization* **8,** 604–616.

Fletcher, R. H., Fletcher, S. W., and Fletcher, G. S. (2012). *Clinical Epidemiology: the Essentials*. Lippincott Williams & Wilkins, New York.

Floudas, C. (2002). Mixed-Integer Nonlinear Optimization. In Pardalos, P. and Resende, M., editors, *Handbook of Applied Optimization*, pages 451–475. Oxford University Press, Oxford.

GAMS Development Corporation (2013a). *GAMS - A User's Guide, GAMS Release 24.2.1*. GAMS Development Corporation, Washington, DC, USA.

GAMS Development Corporation (2013b). *GAMS - The Solver Manuals, GAMS Release 24.2.1*. GAMS Development Corporation, Washington, DC, USA.

Gerald, C. F. and Wheatley, P. O. (1994). *Applied Numerical Analysis*. Addison-Wesley Publishing Comp., New York, 5th. edition.

Griner, P. F., Mayewski, R. J., Mushlin, A. I., and Greenland, P. (1981). Selection and interpretation of diagnostic tests and procedures. Principles and applications. *Annals of*

Internal Medicine **94,** 557–592.

Gupte, M. D., Murthy, B. N., Mahood, K., Meeralakshmi, B., and Nagaraju, B. nad Prabhakaran, R. (2004). Application of lot quality assurance sampling for leprosy elimination monitoring–examination of some critical factors. *International Journal of Epidemiology* **33,** 344–348.

Lemeshow, S. and Taber, S. (1991). Lot quality assurance sampling plans: single and double-sampling plans. *World Health Stat. Q* **44,** 115–132.

Olives, C. and Pagano, M. (2010). Bayes-LQAS: classifying the prevalence of global acute malnutrition. *Emerging Themes in Epidemiology* **7,** 3+.

Olives, C., Pagano, M., Deitchler, M., Hedt, B. L., Egge, K., and Valadez, J. J. (2009). Cluster designs to assess the prevalence of acute malnutrition by lot quality assurance sampling: a validation study by computer simulation. *Journal Royal Stat. Soc.: Ser A* **172,** 495–510.

Olives, C., Valadez, J. J., Brooker, S. J., and Pagano, M. (2012). Multiple category-lot quality assurance sampling: A new classification system with application to schistosomiasis control. *PLoS Neglected Tropical Diseases* **6,** 1–8.

Press, W., Teukolsky, S., Vetterling, W., Flannery, B., and Metcalf, M. (1996). *Numerical Recipes in Fortran 90: Volume 2*. Cambridge University Press, Cambridge.

Robertson, S. E. and Valadez, J. J. (2006). Global review of health care surveys using lot quality assurance sampling (LQAS), 1984-2004. *Social Science & Medicine* **63,** 1648–1660.

Sandiford, P. (1993). Lot quality assurance sampling for monitoring immunization programmes: cost-efficient or quick and dirty? *Health Policy Plan* **8,** 217–223.

Valadez, J. J., Hage, J., and Vargas, W. (2005). Understanding the relationship of maternal health behavior change and intervention strategies in a nicaraguan ngo network. *Social*

Science and Medicine **61,** 1356–1368.

Vanamail, P., Subramanian, S., Srividya, A., Ravi, R., Krishnamoorthy, K., and Das, P. K. (2006). Operational feasibililtiy of lot quality assurance sampling (LQAS) as a tool in routine process monitoring of filariasis control programmes. *Tropical Medicine and International Health* **11,** 1256–1263.

Westerlund, T. and Pettersson, F. (1995). An extended cutting plane method for solving convex MINLP problems. *Computers & Chemical Engineering* **19,** 131–136.

Received October 2016. *Revised February* 2008. *Accepted March* 2008.

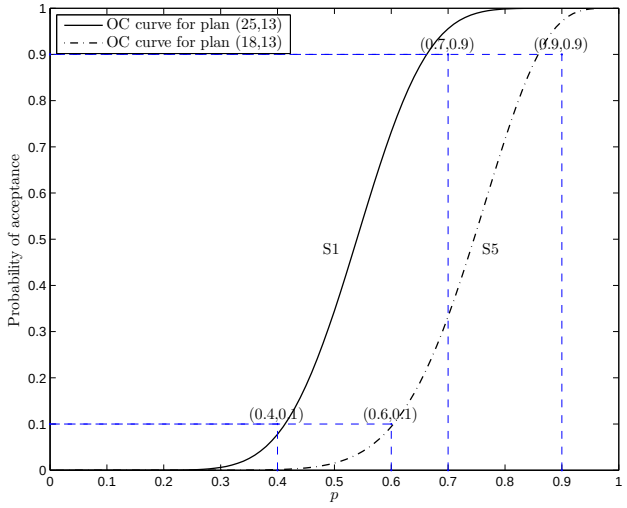

Figure 1. OC curves for LQAS plans obtained for scenarios S1 and S5 employing the *OC curve-constrained design problem* formulation ($\alpha = \beta$ =0.1).

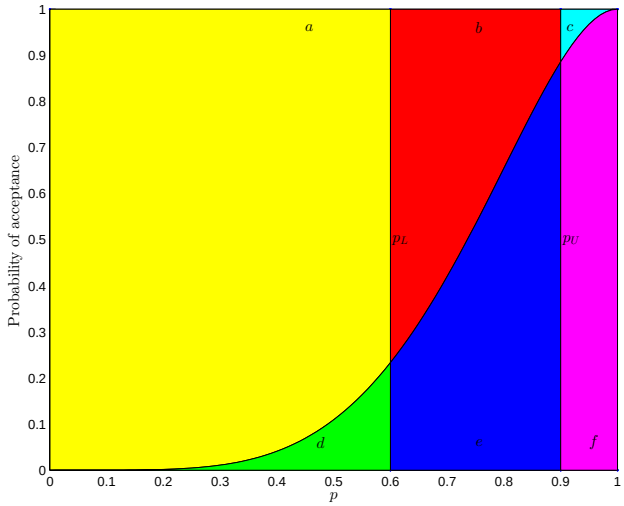

Figure 2. OC curve for LQAS plan obtained with the *performance criteria-constrained design problem* formulation for scenario S5 when the target for Sensitivity for $LAPR$ and for $HAPR$ populations is imposed ($\sigma_U = \sigma_L = 0.95, p_L = 0.6, p_U = 0.9$).

Table 1

Outcomes of the LQAS plan.

		Proportion of individuals with $\xi = 1$ in the population		
		$0 \leqslant p \leqslant p_L$	$p_L \leqslant p \leqslant p_U$	$p_U \leqslant p \leqslant 1.0$
Test Outcome (ξ)	0	a	b	c
	1	d	e	f

Table 2

Diagnosis metrics for different prevalence rates of individuals with $\xi = 1$ in the population.

		Anticipated prevalence	
		Low rate ($p = p_L$)	High rate ($p = p_U$)
Performance metrics	Sens.	$\frac{a}{a+d}$	$\frac{f}{c+f}$
	Spec.	$\frac{e+f}{b+c+e+f}$	$\frac{a+b}{a+b+d+e}$
	PPV	$\frac{a}{a+b+c}$	$\frac{f}{d+e+f}$
	NPV	$\frac{e+f}{d+e+f}$	$\frac{a+b}{a+b+c}$

Table 3
LQAS plans obtained with the OC curve-constrained design problem *formulation ($\alpha = \beta$ =0.1).*

Scenario	p_L	p_U	n	r
S1	0.40	0.70	25	13
S2	0.45	0.75	23	13
S3	0.50	0.80	21	13
S4	0.55	0.85	21	14
S5	0.60	0.90	18	13

Table 4
LQAS plans obtained with the performance criteria-constrained design problem *formulation.*

Combination of criteria	Scenario	p_L	p_U	n	r
Sensitivity for *LAPR* &	S1	0.40	0.70	8	4
Sensitivity for *HAPR*	S2	0.45	0.75	9	5
(Constraints: (10a,10b))	S3	0.50	0.80	7	4
($\sigma_U = \sigma_L = 0.95$)	S4	0.55	0.85	8	5
	S5	0.60	0.90	6	4
Specificity for *LAPR* &	S1	0.40	0.70	12	6
Specificity for *HAPR*	S2	0.45	0.75	13	7
(Constraints: (10c,10d))	S3	0.50	0.80	17	10
($\theta_U = \theta_L = 0.75$)	S4	0.55	0.85	13	8
	S5	0.60	0.90	18	12
Sensitivity for *LAPR* &	S1	0.40	0.70	3	2
Specificity for *HAPR*	S2	0.45	0.75	3	2
(Constraints: (10a,10d))	S3	0.50	0.80	3	2
($\theta_U = 0.75,\ \sigma_L = 0.95$)	S4	0.55	0.85	3	2
	S5	0.60	0.90	4	3
PPV for *LAPR* &	S1	0.40	0.70	4	2
PPV for *HAPR*	S2	0.45	0.75	5	3
(Constraints: (10e,10f))	S3	0.50	0.80	4	3
($\pi_U = \pi_L = 0.6$)	S4	0.55	0.85	5	4
	S5	0.60	0.90	8	7
NPV for *LAPR* &	S1	0.40	0.70	6	3
NPV for *HAPR*	S2	0.45	0.75	7	4
(Constraints: (10g,10h))	S3	0.50	0.80	6	4
($\varrho_U = \varrho_L = 0.95$)	S4	0.55	0.85	7	5
	S5	0.60	0.90	5	4
PPV for *LAPR* &	S1	0.40	0.70	4	2
NPV for *LAPR*	S2	0.45	0.75	5	3
(Constraints: (10e,10g))	S3	0.50	0.80	4	3
($\pi_L = 0.6,\ \varrho_L = 0.95$)	S4	0.55	0.85	5	4
	S5	0.60	0.90	5	4

Table 5
Areas of the OC curve for LQAS plan obtained performance criteria-constrained design problem *formulation for scenario S5 when the target for Sensitivity for* LAPR *and for* LAPR *is imposed* ($\sigma_U = \sigma_L = 0.95, p_L = 0.6, p_U = 0.9$).

						Sensitivity for *LAPR*	Sensitivity for *HAPR*
a	b	c	d	e	f	$\frac{a}{a+d}$	$\frac{f}{c+f}$
0.5733	0.1369	0.0041	0.0266	0.1631	0.0959	0.9557	0.9590