This paper was presented at a meeting of the Republic of Ireland Branch of the Institute on 22nd March, 1971.

A Complete Quality Control System

by A. KRAMER (The University of Maryland, U.S.A.), and T. R. GORMLEY (Kinsealy Research Centre, Dublin, S).

Many people in the food industry feel that if they perform a test somewhere along the line they have performed their duty in relation to quality control and this is all that needs to be done. Any work done in quality control for maintaining a quality level is wasted unless a complete quality control system is used, i.e. that action results from the operation. A complete quality control system consists of a cycle which begins and ends with the buyers requirements or specifications (Fig. 1). The specifications are the heart of the system and the purpose of quality control is to satisfy the buyers specifications at the least cost.

The tests to be done on the product must be established and a sampling procedure decided upon. This is essential since it is usually not possible to test every unit of the product especially if the test is a destructive one. After testing, the results must be reported in such a way that they will lead to action if it is needed.

Points of sampling and testing

Starting with the buyers specifications, samples are taken and tests done at certain locations within the operation. The most important station for sampling is at the point of acceptance of the raw material i.e. supplies, packaging materials etc. There are three good reasons why this is the most important part of the inspection or testing system: (1) this is the only point where a material can be rejected if it is not satisfactory, (2) payment can be in accordance with quality using a sliding scale payment system, (3) a decision can be made as to how the material will be handled in the plant depending on its quality in order to come up with a final product that will satisfy the buyer. Any blending, trimming, sorting and other operations can be done at this stage, as indicated by the test results.

Sampling and testing done subsequently is far less important than that done at the point of acceptance. There may be one or several sampling stations in the operation. These could be at a trimming point, a filling point or at any critical point in the operation, e.g. a station to determine if blanching and retorting is adequate.
Care must be taken not to have too many stations for in-plant inspection and when introducing a new process, new sampling stations and tests should be fully justified in terms of reducing production costs or controlling a critical stage of the operation. In-plant sampling and testing ensures that out of control situations will be avoided. However, it is essential that tests are done immediately after sampling and results reported quickly to ensure that prompt action is taken. In addition time intervals between sampling should be realistic.

The next point at which sampling can be done is on the finished product. This is simply a post-mortem situation because, if through bad control up to now the product does not suit the buyer, then the processor is stuck with something he cannot sell. End-product inspection therefore should be something more than a simple check to ensure that the product is in control. If it has to do more than this it is a sign that control at previous stages was not adequate. In some cases food products in storage or in the channels of trade can also be checked. This again is a post-mortem situation, however, it may be necessary to determine if a product is still marketable after storage for six months or a year.

**Sampling methods**

Sampling procedures fall into two categories, attributes and variables. Attributes plans are generally used by regulatory agencies and unfortunately so, because most of the information obtained is not used. The plan simply tells whether the lot meets specifications or not, so it is a case of accept or reject. Attributes plans are generally used to assess acceptability in relation to the number of defective units but not the severity of the defects. The variables plan works on the basis of actual values obtained for units of the product that have been tested. Payment is made on the basis of these values. If a raw material therefore is of very high quality payment will be high, while payment will be correspondingly low for poor quality.

Since 100% sampling is not usually feasible the sample size required to obtain as accurate an estimate as possible of the true values of the particular lot being tested must be obtained. Generally speaking, the number of samples required for an attributes procedure is the square of the number required for a variables system. Attributes testing requires no prior knowledge of the nature of the material and is based on operating characteristic (OC) curves. These curves relate acceptable quality level (AQL) (abscissa) and the probability of acceptance (ordinate). If a buyer decides that he will not accept more than 1% defective units (AQL=1%) then the probability of acceptance as shown in Fig. 2 is 95%, which means that if a vendor provides material that just barely meets specifications his product will be accepted 95% of the time. The risk of 5% is the vendors risk. In the OC curve (Fig. 2) there is also a buyers risk, e.g., the buyer will accept material containing 2% defects 5% of the time or that with 3% defects 20% of the time (see solid line). Generally these plans are based on a vendor’s risk of 5% of non acceptance and buyers risk of accepting a specified higher defect rate 20% of the time. The larger the number of samples taken the steeper the curve becomes. If the number of samples tested was increased (say from 20 to 200) the curves in Fig. 2 would remain similar at the top i.e. the vendor’s risk would remain the same, however, with a steeper curve the buyers risk would be greatly reduced and he would only accept loads containing about 1½% defects 20% of the time (see dotted line Fig. 2). Tables can be obtained which relate sample size to different risks both for the vendor and buyer.

**Fig. 2** Operating characteristic (OC) curves for an attributes sampling plan

---

161
In variables procedures advance knowledge of the variability of the quality factor of importance in the product is required. For example, in ground beef for hamburger patties, the factor of importance might be the fat content. Knowledge of the fat content of previous consignments is essential for calculating the standard deviation. This quantity is required for use in Equation 1. Calculation of the standard deviation by the elegant statistical procedure is time consuming, however, it can be approximated rapidly by the range method. In this method the range of values (in this case fat contents) among the samples tested is obtained and divided by a d, factor (for the same number of samples) obtained from a table. Equation 1 is used to obtain the number of samples that have to be analysed to obtain the true fat content of the ground beef within certain "precision limits". For example, if it is desired to measure the fat content of a lot to within ±2% of the true value then would have a value of 2 in Equation 1. The number of standard deviations (k) taken give statistical assurance. Taking one standard deviation (k=1) gives the assurance that the fat content will be measured within ±2% of the true value about 66% of the time. A k value of 2 gives about 95% assurance and k=3 gives 99% assurance. It is obvious therefore, from Equation 1 that in order to obtain a high assurance (k=3, 99%) with close precision limits (say ±2%) more samples (n) must be analysed than for a low assurance (k=1) and wider precision limits (say ±5). The buyer therefore has to decide which is more practical for him in terms of cost of analysis and the benefit accruing from a high assurance of being close to the true fat content for a lot. If the cost of doing a fat analysis is small he will probably be justified in testing a number of samples to give him a value very close to the true fat content and thisensure that he is paying only for the fat content he gets. If the cost of testing is high he will test fewer samples and compromise between the lower cost of testing and the fact that he may be paying for a consignment of ground beef that contains much more fat than the testing procedure shows.

**Characteristics of a test for quality**

Tests carried out on a product can be subjective or objective. Subjective tests use human evaluation while objective tests use an instrument or a chemical method. The objective test is preferable in every case and if there is no alternative to a subjective test it is an admission of failure, i.e. there is not a good objective method available. The difference between an objective and subjective test is in a sense paradoxical; after all the response of a human organism to a food is the critical factor and if any instrument or chemical test is to indicate accurately what the human response will be it must be very closely correlated to the human response. The first characteristic of a "test" for quality, therefore, must be accuracy. Subjective tests must, by definition, be accurate. To find out if an instrument or test predicts human response for a certain characteristic in a food, e.g. sweetness, texture etc., the food is tested both objectively and subjectively. A large number of evaluators who are representative of the target population for which the product is intended are normally used and not just a single individual. The results for the samples tested by each method are ranked in order of magnitude and the rank correlation coefficient calculated using Equation 2. This is a rapid approximation method for calculating the correlation coefficient if the correlation coefficient between the two tests is 0.9 or greater it can be taken that the objective test is accurate enough and can replace the subjective one. The coefficient of determination is (0.9)^2=0.81 and shows that 81% of the variation in the subjective test is being explained by the objective one. The remaining 19% is probably error.

\[
\rho (\text{rank}) = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}
\]

- Equation 2

\[
p = \text{rank correlation coefficient} \\
d = \text{difference between the ranks} \\
n = \text{number of samples tested}
\]
The second characteristic of a test for quality is precision, i.e. the ability to duplicate the result on a duplicate sample. Ordinarily an objective test is more precise than a subjective one.

The third aspect of a good test for quality is calibratability. In some cases this can be considered simply as a standardisation of an instrument. If an instrument cannot be calibrated against some absolute term such as lb force, lb/sq. in. or in relation to a particular wavelength or some physical or chemical term it is impossible to know if it is giving the same results on the same sample. This is not so difficult when dealing with one instrument in one location. However, if two or three instruments in two or three locations are being used, results on essentially the same material may be different if the instruments cannot be calibrated against an absolute reference. Alternatively the same results may be obtained with different material. The opportunity for adequate calibration, therefore, is an absolute essential in establishing company-wide, nation-wide or international standards which will be the same wherever they are used.

Tests which may be done

The tests to be performed can be separated into three broad categories. The first is quantitative tests, which as their name suggests relate to quantity such as net and drained weight, underfill, over-run etc. i.e. is the buyer getting the quantity he is paying for. The samples can be weighed on a scale and the test is therefore objective.

The second category of tests deals with hidden aspects of quality such as nutritive value and wholesomeness. Consumers are becoming increasingly aware of the nutritive value of foods and it is likely that food packs in the near future will contain information concerning some of their components, e.g. the protein content of a processed meat product, the vitamin A and C content of orange juice and the fat content of milk. Wholesomeness of foods relates to the microbiology and toxicology of the product. These tests are usually mandatory and must be performed objectively by a recognised microbiological method or chemical test.

Tests involving the sensory area of food quality are the third and most important group that must be considered. The consumer can sense the quality of the product with the eye, i.e. appearance quality, with the muscle sense or sense of feel, i.e. texture and lastly during consumption, flavour. Of the three categories, appearance is by far the most important, texture is next flavour last. Appearance is so important because it is the first human sense involved and in the case of a first purchase it is the only sense that is involved. It is only after the purchase is taken home and eaten that the other factors come into play. A product generally does not stand a chance of selling unless it looks good and this even applies to the container in which it is packed. The reason for demoting flavour to last is that in many foods the flavour of the dish eaten depends so much on the way it is prepared and on the condiments used and not on the intrinsic flavour. Probably the one large exception is in the case of fruits which are eaten largely for their flavour.

Appearance and textural characteristics can be measured objectively, e.g., colour can be measured by a reflectance meter, defects can be counted, texture can be tested with a shearing device and viscosity with a viscometer. In the case of flavour which is sub-divided into taste and aroma, the taste aspect, i.e. sweet, sour, salt and bitter can be measured by instruments or chemical tests. The effective measurement of aroma is much more difficult and has been successful in only a few cases. Normally GLC is used for this purpose. Very often it is off-flavour or off-aroma that is in fact being measured. For example, the peroxide value is an objective index of rancidity.

![Sensory quality circle](image-url)
It should be pointed out at this stage that a certain degree of off-flavour is often desirable and necessary. During storage for 6 months both soya bean and cotton seed oils developed some off-flavour (presumably rancidity) which makes them preferred over freshly refined oils.

There is sometimes confusion concerning clarification of aspects of sensory quality. For example, does consistency fall into category of appearance or texture; are defects classified under appearance or flavour? This can be overcome by representing aspects of sensory quality on a circle as shown in Fig. 3. In this system items such as consistency and defects can be shown as belonging to any two of the major sensory areas represented on the circle.

Reporting results

Results of tests must be reported at once to enable a decision as to whether action is required or not. Results can be recorded in books or flow sheets but undoubtedly one of the best ways of recording is on a control chart. An X variables control chart is based on the standard deviation obtained from tests carried out on the product. Using the example, mentioned already, of the ground beef for the manufacture of hamburgers, the standard deviation for the fat content can be calculated by the range method. If it is assumed that a maximum fat content of 25% is allowed then making patties with a mean fat content of 25% would result in half the patties having a fat content of less than 25% and half with a fat content of more than 25% and hence out of control. Therefore, the mean must be set at less than 25%. If it is assumed, for example, that the standard deviation is 4 then setting the mean at three standard deviation below the upper control limit (UCL) of 25% should ensure that the fat content of hamburgers would only be above 25% one time in a hundred. The control chart, therefore, would show, UCL=25%, mean=13%, LCL=1%. This is a very wide range of fat content and is not satisfactory. However, the range can be narrowed considerably by testing a number of samples and dividing three times the standard deviation by the square root of the number of samples tested. If 9 samples are tested the mean is now $3 \times 4$ or 4% from the UCL, so the chart would now show UCL=25%, mean=21% and LCL=17%. If this is still considered too wide a spread it can be closed still further if it is decided that 95% assurance is good enough. This requires only two standard deviations so with 9 samples the mean falls $(2 \times 4)/3 = 2.67$ below the UCL, i.e. 22.33 and the LCL becomes 19.66%. If these levels are taken as satisfactory then no action is taken provided the fat content remains below 25% and above 19.66%. If the value rises above 25%, less fat must be added during the blending and mixing stage since the product is outside the buyers specifications. If the value is below 19.66% then fat must be added since the product is more than meeting the buyers specifications. Action may also be taken on the basis of trends even though no point is outside the control limits. If six successive values are going in the same direction it is sufficient proof for action. It may indicate a worn bearing on a mixing machine or that the product has settled out or stratified. Six consecutive values above or below the mean also indicates action. The six consecutive values below the mean indicate that the product has become more uniform and the limits should be recalculated thereby giving a saving.

The control chart, therefore, is very much a “visual thing” and should be placed on the operating line where operators can see it. It removes the responsibility of operators having to decide whether action is needed or not on the basis of their own judgement. They simply obey the “rules”, governing the chart and act accordingly. The complete quality control system is now “completed”.