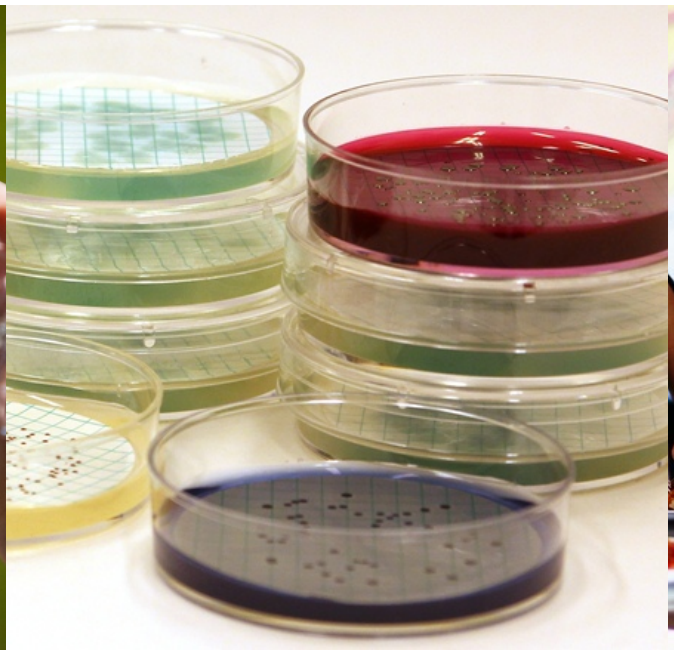
Proficiency Testing

PROTOCOL

Microbiology

Drinking water & Food


Livsmedelsverket
Swedish Food Agency

PT micro
Since 1981

A resolution on accreditation of the program was made 2004-12-10

SWEDAC ACKREDITERING
1457
ISO/IEC 17043

Sixth edition
2022-01-20

Editor-in-chief
Maria Sitell, head of Department of Biology

Proficiency Testing Microbiology	*Webb* www2.slv.se/absint	*E-mail* PT-micro@slv.se	*Phone (operator)* +46 (0)18 17 55 00

Contents

1. Introduction

Laboratories performing analyses need to know that they obtain realistic results. They also need to prove this to their clients to be trustworthy. They can obtain this knowledge by performing different kinds of controls. Internal controls may be used to see that no unexpected changes have occurred within the laboratory. Since microbiological analytical results are dependent on the analytical method used, it is also important for the laboratoriy to be able to compare their analytical results with those of other laboratories. One means of this is to participate in interlaboratory comparative tests. These tests, when they are available, are compulsory for laboratories that aspire to become – or already are – accredited for their analyses. This is e.g. a requirement according to the standard EN ISO/IEC 17025 [1], where the name Proficiency Testing (PT) is used for these interlaboratory comparisons.

PTs are normally organised by a third party, independent of the laboratories and their clients. The evaluations are made by this party and the process is defined as an external assessment of the analytical competence of the laboratory.

The Swedish Food Agency provides PTs in the areas of food microbiology and drinking water microbiology and is accredited for this according to ISO/IEC 17043 [2]. The PTs are mainly directed to accredited laboratories within these analytical areas. They are also suitable for non-accredited laboratories, e.g. laboratories that perform controls in production lines of food or drinking water, but wish to be able to compare their analytical results with other laboratories.

The purpose of this protocol is to give participants, and other interested laboratories and parties, a description of the organisation of these microbiological PT-activities and how some basic tasks are performed. In particular this protocol includes general parts that are not described elsewhere, including the general aspects of production and handling of test material and the statistical processing of results. General information as well as specific information about the testing rounds is available on the program website: www2.slv.se/absint.

2. Organisation

2.1. General information

Address: Swedish Food Agency
Box 622
SE-751 26 Uppsala
Sweden

Telephone: +46 (0)18 17 55 00

The Swedish Food Agency is the central Swedish authority for food issues, including drinking water. Therefore, the Swedish Food Agency organises

microbiological proficiency testings, divided into one scheme for food and one scheme for drinking water.

E-mail for questions and opinions concerning the schemes: PT-micro@slv.se

2.2. Information on the program website

2.2.1. Structure of the website

A general description of the schemes can be found on the website: www2.slv.se/absint.

The website consists of two parts: one public and one for participants only, the latter where a user identity is required to log in.

2.2.2. The public section

This section contains pages that deal with:

- The proficiency testing activity in a general perspective, including the most recent edition of this document as a pdf file.
- Basic information regarding the drinking water and food schemes, e.g. which analytical parameters they comprise.
- The content of current and coming testing rounds along with the time schedule.
- The conditions for participation.
- The current prices.
- Addresses to contact the Swedish Food Agency as provider.
- A web form to apply for registering as a new participant in a scheme.
- The login to participants' pages.

2.2.3. The participant section accessible by login

This section comprises:

- One page to report the analytical results from ongowing rounds.
- One page with web forms, to report details for each method to be stored in our database.
- One page with the reported results and the calculated results from previous rounds that are presented 1) as preliminary shortly after the reporting deadline (to compare your reported results with the preliminary statistics), 2) if necessary: as preliminary during processing of the results, when only the reported results and the preliminary median values are presented, 3) as final when a proficiency testing round has been fully evaluated and the laboratory results and the statistics correspond to those present in the final report. *(Note: as the web page is not as manually flexible as the written report, in cases of discrepancies, the written report overrules the web page.)*
- One page for electronic registration/cancellation of participation in coming rounds.

- One page with the reports and information to the participants (e.g. information letters) available as pdf-files. *(The reports are also publically available on the website: https://www.livsmedelsverket.se/en/PT-micro.)*

2.3. Areas of responsibility

For each scheme, there is a *scheme coordinator* who has the overall responsibility but also specific responsibilities for planning, correspondence, processing of results and reports. For both schemes together, there are *laboratory staff* responsible for the manufacture, quality checks and storage of the test material. There is also an *administrative support* responsible for the participant database, contacts regarding invoices and participation, as well as dispatch of the test items.

The overall responsibility to maintain the quality and the quality system of the program to a necessary level is assigned to the team manager together with the head of the Department of Biology.

The head of the Department of Biology is also the editor-in-chief of the final reports and this scheme protocol.

2.4. Advisory group

2.4.1. Composition and meetings

For the program, there is an advisory group composed of representatives for both drinking water and food microbiology from the countries with the largest numbers of participating laboratories, so far the Nordic countries. The advisory group meets semi-regularly, discussing both issues common to the two schemes and issues specific to the individual schemes.

2.4.2. Duties

The members can come from different organisations and countries, and represent both these and themselves. Their role is mainly advisory, with opinions on e.g. analytical parameters, frequency, costs, accepted methods and the content of the reports. Major changes within a scheme should be sanctioned by the corresponding advisory group. This can be accomplished by e-mail conversation.

2.4.3. Members

A list of the group members is available as a pdf-document under the heading General scheme information under the tab Info & Reports in the participant section of the website.

2.5. Accreditation

The Swedish Food Agency is since December 2004 accredited for arranging microbiological proficiency testing. The accreditation has been approved by Swedac, and is currently performed according to the standard EN ISO/IEC 17043:2010 [2].

3. The two schemes

3.1. The Drinking water scheme

The Drinking water scheme has been ongoing with participants from several countries since 1992.

The scheme comprises 2–4 test items twice a year. It includes about 15 quantitative analytical parameters of bacteria, moulds and yeasts with a focus on indicator organisms, including some that may cause illness.

Some parameters are part of both testing rounds. The remaining parameters are allocated between the two rounds. The parameters are presented on the page General Information, Drinking water scheme, on the website.

3.2. The Food scheme

The Food scheme has been ongoing with participants from several countries since 1988.

The scheme comprises 3–4 test items 3 times a year. It contains more than 25 different quantitative and qualitative analytical parameters – bacteria, moulds and yeasts – including the analyses of pathogenic bacteria normally searched for in food.

Some parameters are part of two or all testing rounds. The remaining parameters are allocated between the three rounds. The parameters are presented on the page General Information, Food scheme, on the website.

3.3. Time schedule and analytical parameters for a testing round

The time schedule and analytical parameters for current testing rounds are found on the page PT rounds on the website.

As a provider, the Swedish Food Agency aims to keep the time frames stated on the website and in this protocol. In case of unforeseen events, we do however reserve the right to reschedule a PT round, or ultimately to cancel it. The participants will be informed about such actions prior to the originally scheduled dispatch date of a round.

4. The test material

4.1. Type of material

Natural samples, or for the particular purpose manufactured test items, may be used in proficiency testing. A variant in between, where a cultured test microorganism is added to a natural or artificial product (”spiking” of a food or drinking water), may also be used at times.

The Swedish Food Agency has chosen to use manufactured test items for the microbiological proficiency testing. These simulate food or drinking water samples containing mixtures of microorganisms, with each test material having a certain purpose.

Depending on the parameters of the round, an individual test item may include both bacteria and fungi (moulds and/or yeasts). Some test items include pathogenic bacteria, while others contain microorganisms possessing specific indicative properties only.

No test material containing protozoa or virus is manufactured.

The ready-to-use test items consist of 0.5 ml freeze-dried serum broth with different microbial mixtures in 2 ml glass vials. The material is manufactured according to the description by Peterz and Steneryd [3]. The sample for testing (simulated water respective food homogenate) is obtained after dissolving the material in a specific volume of suitable liquid.

Freeze-dried test material without matrix in vials has a number of advantages compared to natural samples or "spiked" samples, but also some disadvantages.

4.1.1. Advantages

+ The vials do not require much space, which facilitates storage, package and transport.
+ The vials are fairly robust and are therefore easy to transport.
+ Several organism groups (analytical parameters) may be tested in parallel by the same test item.
+ The test material has a good stability and may be used in more than one PT round without appreciable change of microbial concentrations.
+ It is possible to manufacture a large number of identical sub-samples where the organisms are Poisson distributed at lower concentrations or log-normal distributed after dilutions at higher concentrations.
+ Expenses are kept low by having a standardised and relatively simple procedure of manufacture.
+ Dissolved test material may be used for "spiking" of natural matrices.

4.1.2. Disadvantages

– The freeze-drying procedure and machinery need to be sufficiently stable to ensure that the process can be repeated.
– The organisms require a protective substance, cryoprotectant, which helps them survive the freeze-drying.
– The material has to be dissolved in liquid, which requires a certain amount of work and can induce mistakes.
– The ingredients of the material may cause some foaming when it is dissolved. Therefore, the sample is likely to be distinguished from real samples, which means that the analyst may know it is a control sample.

– For food analyses, natural matrices are absent during preparation of the test material, since the dissolved test material mimics a ready-to-use homogenate.

4.2. Production and product control

4.2.1. Composition

Freeze-dried cultures of the microorganisms used in the test items are stored at −70 °C in a collection (SLV) at the Swedish Food Agency. All strains are identified by specific SLV numbers. The strains have either been isolated from food or water samples, or have been bought from established culture collections. Bacterial strains are characterized internally by API system or by other means at external culture collections like ATCC (American Type Culture Collection) and CCUG (Culture Collection University of Gothenburg). The characterization of yeasts and moulds, *Salmonella* strain and *E. coli* O157 strains is performed externally by, the Centralbureau vor Schimmelcultures (CBS-KNAW Collection, The Netherlands), National Veterinary Institute (SVA, Sweden) and the Public Health Agency of Sweden, respectively.

ISB (Inositol Serum Broth) is used as cryoprotectant to protect the organisms during the freeze-drying process. ISB is composed of sterile-filtered horse serum mixed with inositol and a small amount of nutrient broth. After washing or diluting microorganism cultures from nutrient rich medium (see below), the component SPG (Saccharose Phosphate Glutamate), which also possesses organism-protective characteristics in the freeze-drying process, is added. Peptone water or a potassium phosphate buffer containing magnesium salt (see e.g. ISO 8199:2018 [4]) is used when diluting and washing. The same solutions can be used when dissolving ready-to-use material before analysis.

4.2.2. Manufacture and storage

Each microorganism is cultured individually in suitable nutrient medium, usually TSB (Tryptone Soy Broth) or BHI (Brain Heart Infusion) Broth/Agar for bacteria, and MEA (Malt Extract Agar) for fungi. The cultures are used after a predetermined time period of growth, after which the nutritious medium is washed or diluted off. The test mixture of organisms is obtained by transferring predetermined volumes of each diluted/washed culture (or spore suspension for moulds and other spore formers) to a specific volume of ice-cold ISB.

The test mixture is kept on ice and continuously stirred while aliquots (0.5 ml) are transferred into sterile glass vials. The vials are then freeze-dried and therafter closed by rubber stoppers still under vacuum.

The produced test items are transferred to a freezer (−18 to −24 °C) and tested for its microorganism content shortly after the manufacture. The vials from accepted mixtures are checked for vacuum and are sealed with aluminium caps. The test items are thereafter stored at −55 °C until used. This temperature is chosen to minimize the risk of "glassing", i.e. stiffening, of the rubber stoppers causing air

inlet into the vials, that might occur at a temperatures somewhere below −60 °C. Labelling of the vials is made in connection with the dispatch.

4.2.3. Quality checks

4.2.3.1. Purity of the strains

The purity of all strains included in the test items is controlled by direct culture from the strain collections onto unselective medium. In addition, the purity of each obtained culture is checked by streaking an aliquot on unselective medium. If there is any doubt in the purity of a strain, it is excluded from use and the production of the test mixture is usually postponed.

4.2.3.2. Amount of test material

While dispensing the test mixtures into vials the dispensed volume is monitored by weighing. The total range and measure of dispersion (coefficient of variation) are calculated from the weights. The range may be at most 0.015 g between the dispensed amounts, which corresponds to 3.0 % of the average dispensed amount of material (the target volume 0.5 ml). The weight monitoring is done in parallel to the dispensing, in order to be able to adjust or interrupt the dispensing process, if the weights vary too much.

4.2.3.3. Concentration determination

Concentrations of the included microorganisms are determined after freeze-drying by analysing one or several vials of the test material. For new strains, the concentrations in a test item are often determined before freeze-drying as well. When analyses are performed both before and after freeze-drying, the die off of the individual strains by the freeze-drying process can be determined (the reduction factor).

4.2.3.4. Homogeneity

Before the test material is used, the homogeneity must be checked and accepted. In this first test of homogeneity, duplicate analyses of 10 vials from the whole filling process (stratified sampling) are made with appropriate methods for the parameters to be tested. For the test material to be homogenous, certain criteria regarding variation within and between vials (see below clause 8.2. Determination of concentration and homogeneity) must be fulfilled.

4.2.3.5. Vacuum test

An inert environment is necessary in order to maintain the viability and concentrations of the microorganisms in the test material. To ensure long-time stability of the material, the freeze-dried material therefore needs to stay under vacuum after the vials are sealed and capped. Each individual vial is therefore tested for vacuum before storage, performance tests or delivery. Vials without vacuum are discarded. Normally, very few of the newly produced vials need to be discarded.

4.2.3.6. Stability

The stability of many of the organisms included in the test materials has been investigated for several years in different freeze-dried samples. Each PT test material is however not tested in this sense. A renewed concentration and homogeneity check is done only when used after a longer period of storage (see 8.3.). Knowledge of long-term stability of the test material is based on similar material (reference material) that has been manufactured in the same manner, stored for at least 2 years and tested regularly. When stored frozen at −55 °C, most tested bacteria and fungi have shown stability for at least this period of time. Gram-negative bacteria however, tend to decrease somewhat in colony recovery over time, while Gram-positive bacteria and fungal spores are generally unaffected.

Under storage at approximately 25 °C, no noticeable negative effects on the stability of the test material has been observed for at least 3 weeks, except for the bacterial genus *Campylobacter*. At 44 °C during the same time, the recovery is somewhat affected, especially for Gram-negative bacteria. This tendency is even more noticeable after storage at 60 °C.

Considering these observations, the following guidelines apply:

– If stored at normal freezer temperature (−18 to −24 °C) the content is stable for at least one year, with the observation for Gram-negative bacteria as described above.
– If stored in a refrigerator (5 ± 3 °C), the content is stable for at least a few months, and usually considerably longer.
– If stored at room temperature (not above 25 °C), the content is stable for at least 4 weeks, and usually longer.

4.3. Hazards

4.3.1. Risk of infection

All microorganisms used in the schemes belong to hazard groups 1 and 2, as classified by the the Swedish Work Environment Authority [5]. The Public Health Agency of Sweden (former Swedish Institute for Infectious Disease Control) has made a risk assessment concerning the risks of infection by the microorganisms used in the schemes [6]. In their assessment they concluded that – based on the contents and the way the samples are handled – the risk for illness even after consumption of the contents of one vial is considered extremely small. That assessment is the basis for the stipulation regarding handling and transport made by the Swedish Food Agency [7] (see further under 4.4.3.).

4.3.2. Environmental danger

The test material itself consists of horse serum, inositol and microorganisms, and usually also contains nutrient broth, SPG, peptone, potassium phosphate, and magnesium sulphate. The container consists of glass, rubber, aluminium and has a paper label. Since the container with material does not include any specifically

classed or in any other way potentially dangerous chemical compound, it may be discarded in the common waste management after the microorganisms have been rendered harmless by killing. See below under 4.5.4. Destruction of the material.

4.4. Labelling and transport of test items

4.4.1. Randomisation of test vials to laboratories

4.4.1.1. Numbering of vial labels

Before dispatch to the laboratories participating in a PT, the vials are labelled with individual numbers. The label numbering is made automatically and is saved in our database. The numbers are a combination of the 4-digit, confidential laboratory number and a random sample code after a hyphen (e.g. sample 1000-1, 1000-2 etc.). For each laboratory the sample codes are randomly chosen for the different samples. For example, for one laboratory code no. 1 can be from sample A, but, for another laboratory code no. 1 can be from sample C etc.

4.4.1.2. Selection of vials for a laboratory

Before packaging, the vials of each test sample are carefully shuffled. For each laboratory, a vial is randomly picked and labelled with individual numbers, as described above. This activity is carried out on a separate work bench for each test sample in order to avoid errors.

4.4.2. Package of test vials

The individually labelled vials are placed in a transportation tube containing a shock- and liquid-absorbing material. The transportation tube is put either in a padded envelope or in a safety jar that is put in a cardboard box. A safety data sheet, as well as a delivery note, are also added to the package. The safety data sheet describes the content of the vials and that they can be regarded as freeze-dried artificial food samples. The safety data sheet also describes how the material should be stored and destroyed as well as measures to take in case of leakage and physical contact with it. The package is marked with an address label and, when necessary, a customs declaration and/or other markings necessary for transport.

4.4.3. Transport of test vials

Based on the risk assessment made by the Public Health Agency of Sweden [6] and the stipulation made by the Swedish Food Agency [7], the test vials are packed as described above, and sent via ordinary postal means. In addition, a tracking number or courier service is used for certain laboratory destinations.

4.5. Recommended handling of the test material upon delivery

4.5.1. Storage in connection to dispatch of test material

The material is kept cooled before package and transport (see 4.2.2.).

During packaging and transport, the material is kept at ambient temperature.

The laboratories are recommended to keep the received material in darkness and in a refrigerator (+4 °C) or freezer (−20 °C) until use.

When the test material is to be used shortly after delivery, as in PT rounds, the needs of a long shelf life is relatively small. Storage in refrigerator is therefore satisfactory when the material is received. Storage at room temperature (≤ 25 °C) for up to 4 weeks is also in general not critical. The test material should however always be kept in the dark.

4.5.2. Preparation of samples

The preparation is clearly described to the participating laboratory in the instructions for the PT round, which include pictures with explanatory text (see clause 5.). The procedure states that the test material should be transferred to a measured volume (e.g. 250 or 800 ml) of diluent. The suspension with the test material should then be carefully mixed in order to obtain the sample ready for analysis.

4.5.3. Stability of the prepared samples

Once the freeze-dried material has been reconstituted, the microorganism concentrations cannot be presumed to be stable for longer than about an hour, not even after cooling. The prepared sample should therefore be used for analysis within one hour.

At least some bacterial spores can however be maintained for a considerably longer period of time, e.g. several weeks for *Clostridium perfringens*, to be used as qualitative controls.

4.5.4. Destruction of test material

4.5.4.1. Unopened vials containing test material

Before the material can be discarded in a regular waste bin, the microorganisms need to be killed. This can be done e.g. by autoclaving at 121 °C, for a suffient time to ascertain that the entire content has reached that temperature (e.g. 50 minutes). An alternative is to hand in the vials to a facility specialized in the destruction of infectious material.

4.5.4.2. Opened and used vials

Rubber stoppers and glass vials that contain/has contained test material can be discarded in containers for infectious material, which are to be destroyed by a special facility. The aluminium cap can be discarded along with metal waste.

4.5.4.3. Remains of prepared sample

Remains of the prepared sample should be autoclaved at 121 °C for at least 15 minutes, or treated in any other way ensuring the destruction of the remaining microorganisms, before they are discarded.

5. Instructions for a particular testing round

5.1. Time schedule

Dispatch of test material is made 1–3 weeks ahead of the starting date for analyses in a PT round. The instructions are sent by e-mail latest the day after dispatch, and include information regarding the starting date and the final day to report the results. These dates, as well as other dates for the respective PT rounds, are also stated on the website.

5.2. Analyses

The analytical parameters included in each PT round are stated in the instructions, and can also be found on the website.

5.3. Other information

The instructions sent by e-mail also contain information regarding:

- Preparation of samples.
- Particular conditions concerning the different analyses, such as the dilutions or volumes that should be tested or the methods that can be used.
- How to report method information for each analysis on the participants' pages on the website.
- How to report the analytical results on the participants' pages on the website. A printed results sheet is also enclosed, intended as back-up if there should be a problem with the website.
- That the method information is used to group the analytical results per method in the final reports.

6. A testing round

6.1. Participant activities

6.1.1. Instructions and analyses

Participating laboratories are expected to read the entire instruction and thus to handle the test material according to the recommendations. However, participants should also, as far as possible, perform the analyses in the same manner as for routine samples, taking into account any potential restrictions or addenda stated in the instructions.

6.1.2. Reporting method information

It is compulsory for laboratories to report method information for all analytical parameters. It is also a requirement in order for the laboratory to be able to register analytical results on the website. Method information can only be reported on the website and applies until further notice. It can be entered and adjusted any time after logging in, including after the reporting deadline for analytical results, and between testing rounds. The reported method information should represent what was used to obtain the reported analytical results.

The method information provided by that participants will be used to group the results in order to to distinguish and present method differences in the reports. The method information will normally be used in the way it was stored in the data base at the reporting deadline. Although significant statistical differences may be difficult to prove, trends and possible disparities will be discussed in the report, to assist the laboratories in the interpretation of varying results between participants. All possible groups of a parameter will not be discussed each time.

6.1.3. Reporting analytical results

Normally, results are expressed as integers for drinking water analyses, whereas $\log_{10}$ transformed results are the norm for food analyses. Therefore, the way to report the analytical results differs between the schemes. How to report results is thoroughly described in the instructions for the respective schemes (see 5.3.).

Participants must report results before the stipulated deadline, in order for them to be included and processed in the final report. Before the deadline, results can be reported and changed as many times as desired. After the deadline, changes to the reported data can no longer be made.

6.2. Modifying a reported result

6.2.1. The last day for reporting

Participants must report results and method details within the stated time period.

The Swedish Food Agency will normally remind the participants by e-mail a few days before the reporting deadline. In this e-mail, the reporting deadline is also explicitly stated. However, the final responsibility to report results lies on the individual participant.

6.2.2. Corrections

As a general rule, after the reporting deadline, the only allowed adjustments are those that are due to technical reasons (e.g. computer errors) or due to ambiguities/errors made by the Swedish Food Agency, e.g. due to unclear or incorrect instructions. Corrections are normally accepted only after individual considerations.

Reporting mistakes made by the participating laboratory are thus generally not accepted and therefore not corrected. Such errors include mistakes made when entering the results, results reported for the wrong sample/analysis/dilution,

calculation mistakes and results reported in any other way than described in the instruction, such as use of the wrong numeric scale. The Swedish Food Agency does however reserve the right to – exceptionally – allow corrections even after the reporting deadline, e.g. for special circumstances not covered here.

6.3. Possible sources of errors in a testing round

6.3.1. Registration/cancellation

Laboratories are expected to administer (register or cancel) their participation on the participants' website. This self-administration minimises the risk that laboratories that wish to participate do not receive test material, or the other way around. Laboratories can register or cancel for all available individual PT rounds. They may also register as *subscriber* for all or specific series of rounds. This means that the laboratory is automatically registered for upcoming testing rounds in these series until it changes its status.

Laboratories that for some reason are unable to admister their participation can contact the Swedish Food Agency and ask to be manually registered/unregistered.

The registration deadline for each PT round is shown on the website, and a reminder to login and administer the participation is regularly sent by e-mail by the Swedish Food Agency to (potential) participants.

6.3.2. The test material

During filling of the vials, the amount of dispensed test material is systematically checked by sampling and weighing. After freeze-drying, all vials are checked for vacuum, which is necessary for the survival of the organisms. If more than 4 months will pass since this original vacuum test before the sample dispatch date, a new vacuum test is made on at least 10 % of the remaining vials.

6.3.3. Dispatch and transport

Incorrect labelling of samples for laboratories is a possible risk. As described previously (4.4.1.), the vials are labelled and packaged in a way that mimimises the risk of this happening.

Sending samples to the wrong particpant address is possible, but is avoided by careful and continuous comparison of label adresses, laboratory numbers and sample numbers against each other. Addresses in the participants' database are also updated as soon as changes are received from the participants.

In theory, the testing material could be damaged during transport if it is subjected to very high temperatures or strong x-rays. So far however, this does not appear to have been a problem. Test items have e.g. been transported for long distances to warmer countries, without problem. However, see clause 4.5.1. According to PostNord (the main Swedish postal service) only very low doses of x-rays (<1/100 of the dose for dental x-rays) are used for domestic and international goods at Arlanda (the international airport of Stockholm). Since no general negative effect

from transportation has been noticed, it seems likely that neither temperature, nor the doses of x-rays utilized at airports are a problem for the freeze-dried test items.

6.3.4. Incorrect reporting of analytical results and method information

As described in 6.2., reporting of analytical results is normally made by the participants themselves on the website. The results are saved in the database just as the participant entered them. There is no processing in between that can cause incorrect registration. The only exception to this is when quantitative results larger than 10 are reported for food analyses. Such results are assumed to be mistakenly reported as non-$\log_{10}$ transformed, and a $\log_{10}$ transformation is therefore automatically made. With a few exceptions (6.2.2.), incorrectly reported results are not corrected, but are considered part of the proficiency test.

In some cases, for example when there is trouble with the website or internet connection, a participant can send the results to the organiser via mail or e-mail, for manualy entering into the database. In these cases, the participant will afterwards obtain an e-mail with a copy of the reported results. These results should be checked by the participant, and corrections must be sent to the organiser before the reporting deadline.

Reporting of method information is made on the website (6.1.2.). The method information can be altered at any time, and is expected to be kept up-to-date by the participants. The organiser does not make any changes to the method data, unless it is specifically requested by a participant.

6.3.5. Inaccuracies in the final report

Should a substantial error be found in the final report, the participants will be informed by e-mail. The report is adjusted and a new version is published on the website for reports and information. Less substantial errors or inaccuracies, or minor errors that affect only one or a few individual laboratories may – depending on the circumstances – be corrected directly in an e-mail without publishing a new version of the report. Insignificant errors, e.g. spelling errors that do not affect the report in a meaningful way, do not require neither publication of a new report nor an e-mail.

7. Follow-up of analyses

The Swedish Food Agency does not require – or take any responsibility for – that a follow-up of the results is done by the participants. There is also no requirement for participants to report results from such follow-up analyses. Such demands can only be made by the participant itself or by a third party to which the participant is subordinated, e.g. an accreditation body. Such third parties can demand that the participant maintains a certain quality level and takes measures when the quality is questioned. How the follow-up should be done is therefore the responsibility of the participant itself, or in cooperation with the third party.

The Swedish Food Agency does, however, strive to aid the participants as much as possible in their efforts to understand and correct potential errors.

Z-scores (see below under Statistics) are given in an annex of the reports, as a means for the participants to evaluate their analyses. Z-scores are a good tool to evaluate an analytical parameter over several rounds, e.g. by a control chart.

The Swedish Food Agency also facilitates the follow-up by offering extra vials of the test material to participants that ask for it – for a period of up to two years – provided that the stocks last. Each participant can request one extra vial per test mixture free of charge.

8. Statistics and reporting

8.1. General

The statistical processing in the PT includes the following:

- Numerical checking of the amount, concentrations, homogeneity and stability of the test material.
- Transformation of colony counts (cfu) before calculations, in order to obtain a normal distribution and a uniform variance within the range of results for the respective analysis. In the Food scheme, *log_{10} transformation* is made and in the Drinking water scheme, *square root transformation* is made.
- Identification of deviating analytical results, and assigning them as false positive or false negative results, and as as low or high outliers.
- Compilation of all reported results in a table, along with summary statistics (deviating results excluded) and the number and types of deviating results per analytical parameter.
- Visualisation of the results for each relevant quantitative analysis in a histogram for the respective samples.
- Presentation of statistical evaluation and/or result distribution by method variants per sample for the analysis of a parameter.
- Visualisation of the standardised analytical results (z-scores) of each laboratory in an individual box plot.
- Highlighting outliers and false results in the table containing all the analytical results, as well as summarising the number of these for each laboratory below the respective box plot.

Results that – based on the knowledge of the test material – are obviously erroneous (e.g. undoubtedly false), are excluded without any statistical test prior to the identification of outliers.

8.2. Determination of concentration and homogeneity

8.2.1. Matters common to both schemes

8.2.1.1. General approach

Accredited analytical methods are used when determining concentration and homogeneity of a test material. Methods that are not accredited may be used when a new parameter is tested or in certain special cases. This is noted when done.

Concentrations of the various microorganisms in a test material are determined in part to verify that the material possesses all the desired qualities, and in part to obtain reference values to be used when evaluating the laboratory results. The determination of homogeneity is made based on the microorganism concentrations in vials from a stratified design (directly after manufacture), or by randomly selected vials (subsequent stability tests).

The homogeneity of the freeze-dried test material is tested prior to dispatch to the participants, as packaging and shipment of the vials is not considered to interfere with the homogeneity. Aliquots from several vials are consecutively examined by the same person. The relative variation, both between and within vials, varies depending on which parameter is analysed and on the concentration of the target microorganism. A large variation between vials in the homogeneity test is often correlated with a large variation between the results of the participants as well. However, this is compensated for by the fact that the standard deviation for calculation of z-scores (see 8.6.6.) is not fixed, but is a robust measure based on the results obtained by the participants.

The decisive criterion for homogeneity is in principle the same for the two programs. One non-decisive test (ANOVA) included is based on what was stated in the first edition of an international protocol [8]. It was mainly adapted for quantitative chemical analyses and can therefore not be strictly followed for microbiological analyses. For example, the ANOVA is therfore used only for guidance during the testing, and is completed by other measures more appropriate for microbiology (see below). In later editions of the international protocol [9] and other publications the calculations for determination of homogeneity have changed a bit. The new calculations are not considered applicable for the microbiological activity described here, and are therefore not used.

8.2.1.2. Specific calculations for the evaluation of homogeneity

In a homogeneity test, one or both values from a duplicate analysis of a single vial may sometimes deviate substantially from the remaining values obtained from the other vials. In such a case, the Swedish Food Agency reserves the right to re-evaluate the homogeneity with these values excluded. This will be done if it can reasonably be assumed that the divergent results are not due to non-homogeneity of the test mixture, and instead due to e.g. a pipetting error or the analysis of the wrong dilution or volume. If the results from re-calculations with such values excluded fulfil the citeria for homogeneity, the test mixture will be approved.

1 – ANOVA

One-way analysis of variance (ANOVA) is performed on the results from the 10 vials with duplicate determinations [8, also described in reference 10]. It is included here since it is often used in chemistry proficiency testing, where the results often have a normal distribution. The ANOVA is carried out with *square root transformed results* for drinking water and log_{10} *transformed results* for food. These tranformations are made to obtain uniform variances and to achive an as good normal distribution as possible. An F-test is carried out to ascertain that the dispersion between vials is not markedly larger than within the vials.

2 – "Index of dispersion" – check of randomness

As a complement to ANOVA, the "Index of dispersion" test is used. It is a test more suitable for microbiology, and is used to check that the obtained analytical results both within vials (I_1; 10 duplicate analyses) and between vials (I_2; 10 vials) do not differ markedly from what would be expected based on the appropriate Poisson distributions [11, 12]. In these tests, the original *non-transformed colony counts* from the sample volume chosen for counting, in the analysis of the parameter, are used. The test is dependent on concentration in the sense that it is easier to obtain acceptance for randomness, i.e. no contradiction to the Poisson distribution, at low colony counts compared to at high.

3 – Test of reproducibility

This test [13] is independent of the other and of the microorganism concentration. It is therefore a necessary complement to the "Index of dispersion". This test is performed on the 10 average values or sums (both give the same results) from the two results obtained for each vial, expressed in log_{10} *units* to normalize the results in a general way.

8.2.2. The Drinking water scheme

8.2.2.1. Prerequisites and assumptions

Quantitative analytical results, concentration determinations as well as the participant's results, are obtained by manual counting of colonies. A result may be derived from one or more different volumes of a sample (dissolved test material) but is in general converted into a pre-set analytical volume of the sample. Colony counts from different aliquots of the undiluted sample are generally considered to be Poisson distributed. Since dilutions are never made, the Poisson distribution is regarded to be approximately valid, even for the converted results. Strictly however, it is only valid for the test volumes that have been analysed. In a Poisson distribution, the variance is numerically equal to the average number of colonies.

As a consequence of assuming a valid Poisson distribution, a square root transformation is, based on theory [12], used to approximate a normal distribution. This is supported by studies from the early years of the scheme, where in most cases the participants' results had the best normal distribution when they were square root transformed, compared to log-transformed or not transformed at all.

The analytical results are used square root transformed in determinations of concentrations, in the ANOVA and in detection of outliers, where a reasonable normal distribution is a prerequisite. In tests based on Poisson-distributed results, the results are instead used without transformation.

Depending on the testing round and the analytical parameter, the number of quantitative analytical results varies from approximately 40 to 100.

8.2.2.2. Initial control of concentration

As close as possible to the freeze-drying of a test mixture, 5 vials, randomly picked at different stages of the filling process (beginning, middle and end), are analysed. The content of each vial is dissolved in a fixed volume of diluent, and single analyses are performed from relevant sample volumes. This initial check of concentration is done to get an early indication on whether the test mixture appears to be acceptable, and if it can be assumed to be homogeneous with respect to the various organisms included. It is also done to decide which sample volume to use for each parameter in the final determination of concentration and homogeneity.

8.2.2.3. Final determination of concentration and homogeneity test

Before dispatch of the samples, 10 vials per test mixture from the various stages of dispensation process (stratified sampling) are tested with respect to homogeneity and concentrations. The test items are prepared in the same was as for the initial concentration check. Each sample is analysed in duplicate, and the analyses are performed on all 10 vials during the same day under repeatability conditions. The duplicate aliquotes from the 10 vials cannot be analysed in a random order, since unwanted changes in the concentration of organisms could occur in the sample suspension if too much time elapses before analysis. The two aliquotes from the same vial are therefore analysed within a relatively short period of time, first aliquote one on the various media and thereafter aliquote two. This could potentially lead to under-dispersion. Only one specified sample volume, based on the initial concentration check, is used for each parameter. A global average and a coefficient of variation (CV; see 8.6.7.) are calculated from the 10 mean values (each from 2 determinations) per analysis.

8.2.2.4. Criteria for homogeneity

The results from the three tests described in 8.2.1.2. complement each other, and together give a complete picture of the outcome. However they must be interpreted with care and caution, and the results from previous similar samples need to be considered. The calculations from ANOVA and "Index of dispersion" within vials (I_1) are used to understand the contribution of randomness in the outcome of the tests. Two empirically verified indicative limits from the other tests are used as decision makers.

The actual criterion for homogeneity is that the values obtained for the test of reproducibility (T) and the test "Index of dispersion" between vials (I_2) *do not simultaneously exceed the set limits*, i.e. 2 both for T and I_2.

For production of test items before 2015, the guideline for homogeneity to be acceptable was that the coefficient of variation (CV, see 8.6.7.) should not exceed 25 % when the average content was at least 10 cfu per analysed

sample volume. With a lower cfu average (<10 cfu; often poor normal distribution even after transformation), a CV higher than 25 % was accepted if the distribution of the colony counts was as expected, based on the interpretation of the test "Index of dispersion" (with 95 % confidence).

8.2.3. The Food scheme

8.2.3.1. Prerequisites and assumptions

Quantitative analytical results are obtained by manual counting of colonies. These colony counts are thereafter in general converted to a pre-set analytical volume, taking into account the dilution factors. Colony counts obtained from aliquots of the same dilution in a dilution series are in general considered to be Poisson distributed. Since colony counts from several dilutions are often recorded, and since conversion is made back to a specified volume or weight of the original sample, results cannot fully be considered to strictly follow a Poisson distribution. The variance increases more than what is assumed in the Poisson distribution and therefore a log-normal distribution is used as an approximation. The usage of common logarithms ($\log_{10}$) is also the practice when it comes to microbiological food analyses.

The final calculated analytical results are therefore expressed as $\log_{10}$ results and are considered more or less normally distributed. These log-transformed results are used when calculations of concentrations, assessment of homogeneity (except for I_1 and I_2, where original counts are used) and detection of outliers are made.

Depending on the testing round and the analytical parameter, the number of quantitative analytical results varies from approximately 40 to about 200 (in some cases also <40).

8.2.3.2. Initial check of concentrations

After the freeze-drying of a test mixture, a single vial is tested to determine the organism composition and concentrations. The content of the vial is dissolved in a fixed volume of diluent, which is then referred to as the zero dilution. Single analyses are made from various dilutions. The initial check of concentration is used to obtain an indication on whether the test mixture is acceptable regarding the different included organisms, and also to decide which dilution to use for each analysis in the final determination of concentration and test of homogeneity.

8.2.3.3. Final determination of concentration and homogeneity test, quantitative analysis

Before dispatch of a test material, 10 vials from the various stages of dispensation process (stratified sampling) are analysed by one person on the same occasion, and under repeatability conditions. The test material is prepared in the same way as the initial concentration control, and the most suitable dilution to analyse is based on the results from the initial concentration check. The 10 samples are analysed either manually or with a spiral plating equipment. Two dilution series are made from each vial, and one Petri dish per series is used for each analytical parameter. The two dilution series from each vial are not analysed in a random order, since unwanted changes in the microorganism concentrations could occur in the

dissolved and diluted sample if too much time elapses before analysis. The two dilution series from a vial are instead analysed within a relatively short period of time; first all media from one series and therafter all media from the second series. This could potentially lead to some under-dispersion. An average result is calculated for the two Petri dishes from each vial. The mean value of these 10 average results *log_{10} transformed*, and converted back to the original volume using the dilution factor, constitutes the final concentration. The standard deviation for these 10 results, expressed as log_{10} values, is also calculated.

8.2.3.4. Final determination of concentration and homogeneity test, qualitative analysis

To determine the concentrations of target organisms in *qualitative* analyses (e.g. *Salmonella*), separate vials containing a pure culture, "special vials", are freeze-dried at the same time as the vials containing the test mixture. The volume ratios used in the manufacture of the “special vials” are identical to those in the test mixture; the same final concentration of microorganisms are therefore expected in both. Initial control of concentrations, as well as final determination of concentration and test for homogeneity, is performed on 1 and 10 vials, respectively, in the same way as for the test mixture. The mean value per vial is therefore considered to reflect the organism concentration in the test mixture.

Target organisms for qualitative analyses may also at times be quantified directly from the test mixture by direct inoculation on selective agar media. However, due to stress and competition, this may lead to a lower colony recovery than the actual concentration in the vials. In such a case the dispersion will also usually become larger.

8.2.3.5. Criteria for homogeneity

The criteria for homogeneity follow the same reasoning and principle as for the drinking water samples (8.2.2.4.). For food analyses, the criterion for homogeneity is that the values obtained for the test of reproducibility (T) and the test “Index of dispersion” between vials (I_2) *does not simultaneously* exceed the set limits of 2.6 for T and 2.0 for I_2.

For qualitative analyses, the same criteria prevail and the target organisms must be detected from all the 10 vials containing the mixed cultures.

For production of test items before 2015, the criteria of homogeneity for quantitative analyses according to Peterz, 1992 [14] was used. The range for the *log_{10} transformed* values of the mean result of the 10 vials should not exceed 0.5 log_{10} units and the standard deviation should be <0.15 log_{10} units. In the original test of Peterz results from only 5 vials were used.

A value of 2 for T was desirable when T was decided to be used as a measure in the Netherlands [13], but in practice a value of T = 3 were used. Here the value T = 2.6 is chosen for food analyses because it corresponds exactly to the standard deviation of 0.15 log_{10} units that was previously used. For drinking water analyses the desired value T = 2.0 is used based on empirical checks. This difference is justified by the fact that for food analyses more divergent high values are obtained (in theory) that become normalized after log_{10} transformations in comparison with results from drinking water analyses which are normalized by square-roots.

8.2.4. Measurement uncertainty for the final determination of concentration and homogeneity

When a manufactured material is assessed as homogenous, in the sense that the dispersion is not significantly larger between vials than within vials, the individual analyses are regarded as 10×2 (20) *independent analyses*. The pooled standard deviation from the ANOVA is therefore used as the measurement uncertainty for the determination of the concentration that was performed under repeatability conditions during the test of homogeneity.

In these determinations, transformed analytical results are used in the same way as in the ANOVAs, i.e. *square root transformed* values are used for drinking water analyses and *common logarithm (log_{10})* values are used for food analyses.

8.3. Stability

Not all test mixtures with freeze-dried microorganism are tested for stability. However, several weeks often pass between the initial concentration check and the homogeneity tests. To a certain degree, this facilitates the detection of unexpected changes in the concentrations (e.g. due to loss of vacuum in the vials). Mainly however, the Swedish Food Agency relies on decades of experience from working with freeze-dried materials, including continuous follow-ups for several years of a large number of similarily manufactured reference materials (RM)[*].

Some test mixtures are produced in larger batches and used for more than one PT round. Before vials from such test mixtures are used in a subsequent PT round, a new combined stability check of the concentration and homogeneity is performed if more than 6 months has passed since the last check. For such a stability control, 5 randomly selected vials are analysed in the same was a stability test of a RM [**]. If the results from these 5 vials do not fulfil the homogenieity criteria for one or more parameters, the analyses of those parameters are repeated on 10 new vials. If the critera for homogeneity are still not met after those 10 vials have been analysed, the test mixture is not used.

[*] Stability tests for RM have been carried out on a regular basis during at least 2 years for materials stored at −55 °C (−65 °C before 2012) and in many cases at −20 °C. Analyses have been performed after about 1, 6, 12 18 and 24 months after production. In addition, in doubtful cases, the material has been further analysed. Follow-ups are often made in stable samples mixture for up to 36 and 48 months. In order to detect changes better, the results are plotted over time for each analytical parameter and reference material for the temperature it was kept at.

[**] The stability tests correspond to half a homogeneity test, i.e. the content of 5 vials is analysed in duplicate (two equal volumes or two dilution series for drinking water and food, respectively). The statistical analyses correspond to those in the homogeneity tests but are based on 5×2 results instead of 10×2 results.

8.4. Outliers in a testing round

Outliers are results that differ so much from the other results that they cannot be explained by the ordinary variation. Outliers can be objectively identified in different ways. In both the Food and Drinking water schemes, the Grubbs test modified by Kelly [15, 16] is used. The level of 1 % is set as the risk to erroneously

identify a result as an outlier. A prerequisite for a correct test is that the results are normally distributed.

In order to begin with the best possible normal distribution, the results are always transformed before processing. Square root transformation is used for drinking water results and $\log_{10}$ transformation is used for food results. The Grubbs test is used as an objective tool to identify deviating results. This test is used even when the transformed results are not fully normal distributed and, thus, the assumption of normal distribution is not fulfilled.

Outliers are excluded before the final calculations of medians, mean values and measures of dispersion for the various analytical parameters. However, z-scores (see 8.6.9.) are calculated also for the outliers, using the same mean value and standard deviations for a parameter as for ordinary z-scores.

8.5. False positive and false negative results in a testing round

8.5.1. The number of false results

The number of reported false results varies a lot depending on which parameter is analysed, the composition of the sample and the degree of difficulty, e.g. concentration and/or background flora.

8.5.2. False positives – definition

A false positive result is an analytical result where an organism/analyte is reported as detected even though it was not present in the sample.

8.5.3. False negatives – definition

A false negative result is an analytical result where the target organism is reported as not detected in a relevant volume, even though it should be present in that volume.

8.5.3.1. Additional considerations for the drinking water scheme

When the average concentration of a target organism is high, a zero result is obviously false and no outlier test is necessary to confirm this.

When the average concentration of a target organism is not high but not very low, zero results that are identified as outliers are defined as false negative results. Zero results that are not identified as outliers when the average for the parameter is >15 cfu/ml are considered as potentially false negative. In the annex in the final report they receive the remark "False negative?" (underlined).

When the average concentration of a target organism is very low (most often <10 cfu/ml), zero results are often obtained randomly. In these cases they are usually not identified as outliers and, thus, not as false negative results.

8.6. Statistics for a testing round

8.6.1. Transformations and account for different measures

Calculations of mean values, standard deviations and z-scores are carried out on log_{10} *transformed* or *square root transformed* results for the food and drinking water analyses, respectively (see 8.1.).

For drinking water, the mean and median values are given in the normal cfu scale (back transformed) in reports, while the common logarithmic results are given there for the food scheme.

The various measures in this chapter, which are based on normal distributions (all except median and range), are in general not given in the report when there are fewer than 20 results in total for an analysis. In instances when they are still shown, despite fewer than 20 participants, this is explicitly noted in the report.

For analyses with fewer than 20 results, only the median and the range are normally shown. In these cases, false results are reported but not outliers.

The median and range are also shown for parameters that are not assessed or given z scores, and used for discussions only. In these cases, neither false results nor outliers are reported.

In general, when results are grouped based on the method used, measures of dispersion and mean values are only displayed for method groups with 5 or more results. Sometimes they are yet calculated and shown for discussion.

8.6.2. Median

For the preliminary report on the website, medians are provided instead of mean values as the outliers are also only preliminary. Medians are also provided in parallel to the mean values in the final written report. The median is more robust than the mean value, which means it is less affected by outliers and the distribution.

8.6.3. Range

The range is the interval that includes the lowest to the highest reported result for a parameter. Depending on whether the analysis is assessed or not, the interval is used after or before the outliers and false results are excluded, respectively. When deviating results are excluded, the range can be called an *interval of acceptance*.

8.6.4. Mean value

The mean value is calculated as a consensus value from the results of the participating laboratories after outliers and false results have been excluded.

8.6.5. Assigned value (m)

For evaluated parameters, the assigned value is the mean value (8.6.4.) of the participants' transformed results, the consensus value. It is regarded as the true, normative value. In some cases different assigned values may be relevant, based on groupings of the data according to the methods used. For parameters that are not statistically evaluated, the median is stated as the assigned value.

In the drinking water scheme, the re-transformed (squared) assigned value is reported as the mean value rather than the assigned value itself.

There are a couple of reasons for choosing the consensus value as the assigned value, rather than a value determined by expert laboratories.

a. Microbiological quantitative results are strongly dependent on the method, and therefore a “true value” strictly does not exist. Rather, the true value is to some extent a question of definition. Results obtained with a certain method by one or several expert laboratories are not necessarily more correct than those obtained with another method by a participating laboratory.
b. Different brands or batches of dehydrated culturing media – manufactured according to a relevant method standard, and properly evaluated – may give different colony appearances and recovery. A systematic recovery bias of a particular culturing medium should not be allowed to have an impact on the outcome for participants using another medium.

8.6.6. Standard deviation for proficiency assessment (s)

The variation around the participants mean value for an analysis is estimated from the actual variation of their results. It consists of the standard deviation of the results, with outliers and false results excluded. This consensus standard deviation is also used as the denominator in the calculation of z-scores (see 8.6.9.).

The standard deviation is also used as a measure of dispersion in the Food scheme since it is a *relative measure* (independent of the concentration) *when logarithms are used*.

An alternative to this standard deviation, which usually varies from round to round for a parameter, would be to have a fixed standard deviation for each parameter. One way is to use an average value based on previous suitable testing rounds. With such a procedure – as is often recommended – z-sores are considered to be directly comparable from one round to another. However a prerequisite for this to be relevant, is that the test materials of different testing rounds have the same degree of difficulty. If they do not, a more difficult test mixture – with a larger dispersion of the participants’ results – would consequently have a larger number of z-scores outside the acceptance limits. As this would in essense be caused by the manufacturer of the test material, it can be argued it is not something that should affect the outcome for an individual laboratory.

The difficulty to obtain a correct result varies with the test material, both within and between testing rounds. This is true even for the same analytical parameter. The Swedish Food Agency has therefore chosen to use a standard deviation that varies with the difficulty of the test material, i.e. the actual standard deviation for the current test material. In general, this leads to about the same proportion of “high” z-scores for results within the accepted limits (i.e. without outliers) in all testing rounds. Added to this are a varying number of extreme "z-scores" for the outliers (that are not included in the calculations of mean value and standard deviation). The total number of extreme z-scores may therefore vary more or less independently between rounds.

In this way the z-scores are more accurate for the performance over time as well, since a compensation for the degree of difficulty is made. Use of the participants’ standard deviations is therefore considered to be the most appropriate way to make comparisons of the microbiological analyses within the frame of the program.

8.6.7. Coefficient of variation (CV)

The coefficient of variation (CV) is a *relative measure* and is the ratio of the standard deviation (8.6.6.) to the mean, expressed as a percentage.

The CV for the *square root transformed* results is stated as the measure for dispersion in the Drinking water scheme.

8.6.8. Measurement uncertainty for the assigned value

The uncertainty of measurement for an assigned value is calculated as the ”standard error”, i.e. the standard deviation for the parameter divided by the square root of the number of accepted results.

As the number of reported results included in the calculations increase, the influence of the measurement uncertainty decreases compared to the standard deviation for the calculation of the z-scores (see 8.6.6.). Even with only 10 results, the measurement uncertainty constitutes only about 30 % of the standard deviation; it decreases to less than 10 % with more than 100 results.

Accuracy – consisting of trueness (bias) and precision (dispersion) – is not provided for the results. Differences in recovery due to usage of different methods and/or media are a problematic component of microbiological analyses. This means that systematic deviations from a "true" value cannot be quantified and are not necessarily erroneous. The accuracy is therefore only specified as precision and expressed by the measurement uncertainty.

8.6.9. Comparative values for follow-up (z-scores)

All results – including outliers but excluding false results – from the assessed parameters are transformed into standard values (z-scores) according to the formula:

$$z = \frac{x - m}{s}$$

x = the result of the individual laboratory (in transformed format)
m = the mean of the participating laboratories (the assigned value)
s = the standard deviation of the laboratories results around m

After this transformation, the standard values, apart from those based on outliers, have a mean value equal to zero (0) and a standard deviation equal to one (1). They also have a distribution that can be compared to a standardised normal distribution. The z-scores facilitate comparison of the various analyses with each other, since they are independent of the concentration and expressed on the same scale (the number of standard deviations).

In 95 % of the cases, an individual z-score, apart from those based on outliers, will be found in the interval [–2; +2]. The probability to fall outside these limits is less than 5 %. The probability to fall outside the interval [–3; +3] is less than 0.3 %.

Z-scores that are based on outliers can be regarded as artificial, not fully true, z-scores. This is since outliers are not included in the calculations of the consensus mean value and the consensus standard deviation for an analysis. These "artificial" z-scores exist in addition to the ordinary ones and can in general be found outside of the interval [–3; +3].

The z-scores are the base of the box plots (see 8.7.2.4.).

8.6.10. Interpretation of the z-scores

In evaluations of the analytical results, the following guidelines can be used:

$\|z\| \leq 2$	indicates that (the original) result is acceptable
$2 < \|z\| \leq 3$	indicates a warning that the result may be deviating, and might motivate an action in the follow-up process
$\|z\| > 3$	indicates that the result is regarded as deviating and should lead to an action in the follow-up process.

8.7. Reports from a testing round

8.7.1. Preliminary report on the website

Within two weeks after the results reporting deadline, a preliminary report is published in two parts on the participants' pages on the website (see 2.2.3.).

The first part consists of the results reported by the laboratory, together with the preliminary processed results of all participants with median values and a preliminary relative measure of dispersion. The preliminary limits of acceptance are available for at least 2 weeks. As the preliminary limits of acceptance are subject to change, they are not always shown during the final processing of the results. During this period, participants can only see their own results and the medians, until the final statistics and the final report are published.

The second part is a pdf document with comments, which should be read alongside the preliminary results. It contains a short description of the sample content, and a description of how to interpret the results with respect to the preliminary limits of acceptance. The report may also contain an annex with photos that show the outcome and appearance of the colonies on various selective media.

8.7.2. Final report

8.7.2.1. General

A final report is published within 2 months after the results reporting deadline. The report contains comments and discussions on the results for the different parameters, as well as general discussions on the outcome and performances of the parameters and methods. Special attention is given to deviating results, and to instances where an analytical parameter was difficult or impossible to evaluate in a certain sample. The results for each relevant parameter are presented in two histograms (see 8.7.2.3.), one visualising the accepted and deviating reults, and one visualising the different methods that were used. The results are also displayed in tables that show the outcome from the different methods that were used. Possible discrepancies between methods are discussed. An annex with the results of all participants (with identification code) is provided at the end of the report. The table highlights outliers and false results, and contains summarising statistics at the end. Another annex contains the calculated z-scores for each participant, and is intended to assist with the follow-up for the individual laboratories. Z-scores are also graphically summarised in box plots, one per laboratory, in order to visualise the overall performance (see 8.7.2.4.). Finally, the report may contain an annex with photos that show the outcome on various selective media.

The report is available as a pdf-document on the participants' pages of the program website (see 2.2.3.). It is also available on the public website of the Swedish Food Agency (https://www.livsmedelsverket.se/en/pt-micro).

8.7.2.2. Quality control results

For both schemes, results from the final determination of concentration and the homogeneity test (see 8.2.2.3. and 8.2.3.3.) – or the last stability-homogeneity check (see 8.3.) – are presented in a table in the final report.

8.7.2.3. Histograms

A histogram is made for each analysis where quantitative non-zero results are present. It shows the distribution of the results with outliers and false negative results highlighted. When relevant, a corresponding histogram is shown with the results from different methods or media highlighted in different colours. The histograms for drinking water are based on the original colony counts and the histograms for food are based on $\log_{10}$ results. Examples of histograms are shown in figures 1 and 2. An asterisk indicates that the results are outside the adjacent horizontal axis limit.

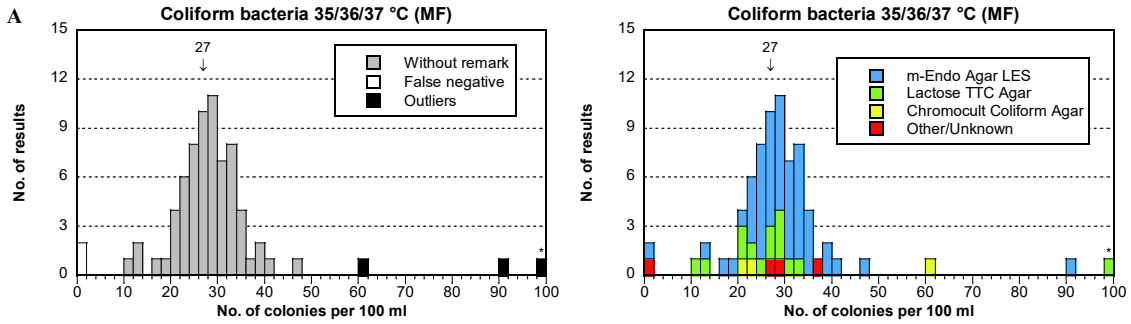


Figure 1. *Examples of histograms for a drinking water analysis.*

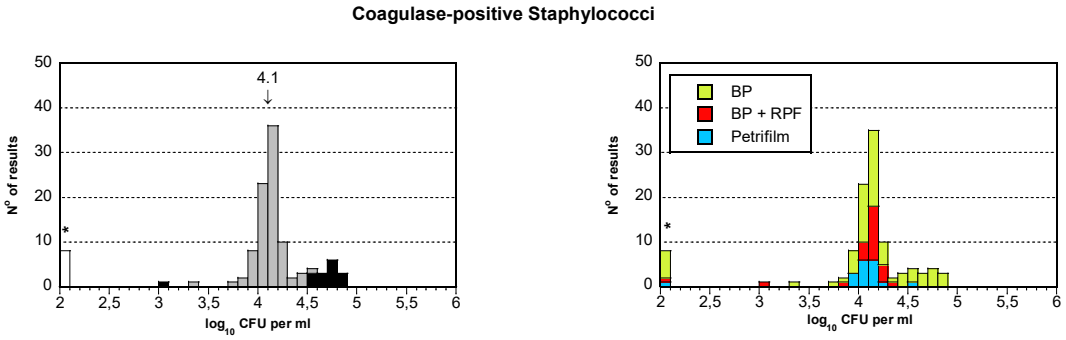


Figure 2. *Examples of histograms for a food analysis.*

8.7.2.4. Box plots

Each box plot is based on the z-scores of an individual participant. It illustrates how the participant's z-scores are situated as a group in relation to the common, "true", mean value zero. The median value of the participant's z-scores is illustrated by a solid horizontal line in the box. The box itself consists of 25 % of the z-scores above

the median and 25 % of the z-scores below the median. The remaining upper 25 % and lower 25 % of the results are represented by vertical lines and circles. A z-score is represented by a circle when it is deviating by a certain degree* in relation to the extent of the box.

* < [lowest value in the box −1.5 × (highest value in the box− lowest value in the box)] or
> [highest value in the box + 1.5 × (highest value in the box − lowest value in the box)]

Z-scores higher than +4 or lower than –4 are positioned at the limits +4 and –4 of the y-axis of the plots, respectively. An example of a group of twenty box plots is presented in figure 3.

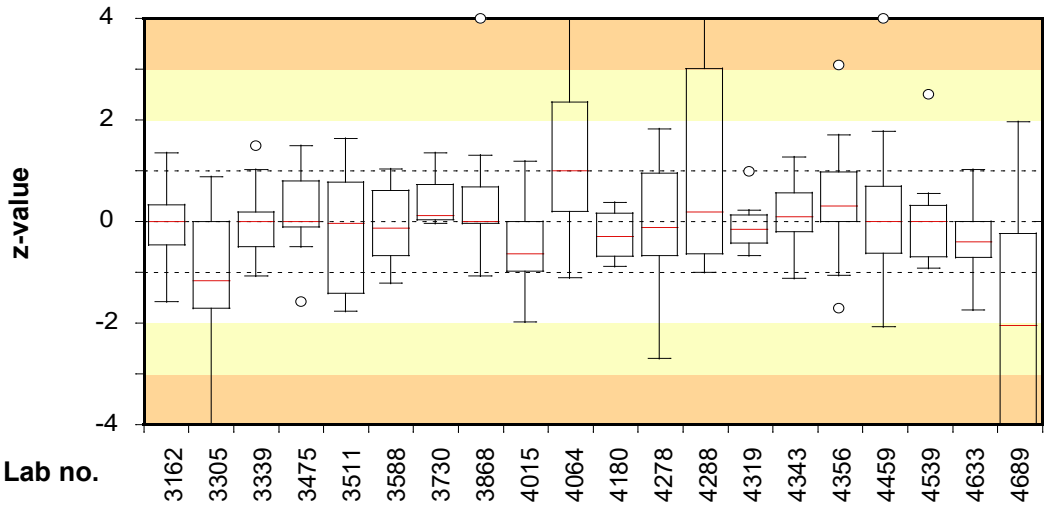


Figure 3. *Example of a graph with twenty box plots.*

8.8. Assessing the performance of an individual laboratory

The performance of an individual participant is not directly assessed in the final report. The basic criteria for assessment are however provided. These are foremost the number of false negative and false positive results, together with the number of outliers (see 8.4. and 8.5.) or the numbers of deviating z-scores. For each participant, the z-scores are presented in the box plot and the other numbers are presented in a table below it (see 8.7.2.4.). The corresponding results are highlighted (bold and shaded area) in the annex of the final report containing the results of all laboratories, making them easy to connect to a specific laboratory. In the same annex, sample numbers (1, 2, 3 etc.) are crossed over if, based on many analytical results, it is obvious, that samples have been mixed-up.

In the final report, circle diagrams that provide a summary illustration of the overall performance of the participants are shown for each sample. Individual participants are not ranked based on their performance; it is thus up to each participant to follow-up and interpret their outcome and performance. The z-scores (see 8.6.9.) and box

plots (see 8.7.2.4.) of a laboratory can be used as an aid to this. Outliers are reflected by high positive (+) or negative (–) z-scores, while false results do not generate z-scores at all. Z-scores are otherwise mainly intended to compare the results of different analytical parameters within rounds, and the results of a specific parameter between rounds.

The box plot combined with the numbers of false results and outliers in the table below the plot is the easiest way for an accreditation body, or other interested parties, to get an overview of the performance of an individual laboratory in a PT round.

9. Confidentiality and user identity

9.1. Confidential laboratory number

Each participant is given a unique laboratory number when registering. The number is valid for the schemes the participant is registered for and is confidential. It means that the Swedish Food Agency reveals the number to no one but the participant in question. The Swedish Food Agency never states the laboratory number to a third party, except when given permission to do so by the participant.

9.2. Confidential password

In addition to the laboratory number, all participants are also given a confidential password when registering.

9.3. Usage of laboratory number and password

The participant shall state its laboratory number when communicating with the Swedish Food Agency regarding a scheme. This number is also used by the Swedish Food Agency to identify the participant on the website, in compilations and in the final reports. Correspondence with a participant where its identity and laboratory number is revealed is treated as confidential.

The laboratory number and the password have to be used together to log on to the participant pages of the website (see 2.2.3.).

The laboratory number and password are stated on a document label enclosed with the test items in the package for a round.

9.4. Changing the laboratory number and password

The laboratory number may be changed in order to minimise the risk of unwarranted usage, e.g. upon staff turnover. The password and the laboratory number shall be changed upon a participants written request, or if either part has used them in such a way that the identity of the laboratory has been revealed.

10. Filing

All results that are reported in the test rounds are filed in the Swedish Food Agency's participant database, for at least 4 years after publication of the final report.

In general, all documents derived from PT participantion, e.g. correspondence related to or generated by PT participation, are filed for a period of at least 1 year.

11. Comments from participants and remarks

11.1. Policy

Remarks in the form of deviations and complaints on the work performed within the PT schemes are documented and investigated. If required, corrective actions and measures to avoid re-occurrence will be taken. Comments and suggestions for improvements are taken into consideration and dealt with in an appropriate way.

11.2. Remarks

Remarks within the PT schemes are divided into *complaints* and *deviations*.

Complaint: when a participant or other interested party is dissatisfied with a service or a product.

Deviation: when written procedures are not followed or when requirements or agreements are not fulfilled.

These definitions and the procedures for managing the complaints and deviations are described in general terms in a general instruction [17] and more specifically in the internal procedure of the program.

Complaints and deviations, as well as the corrective and preventive measures taken, are documented in a database according to internal procedures at the Swedish Food Agency. If the issue is more general it will also be managed within the case and document management system at the Swedish Food Agency.

11.3. Suggestions for improvement

Suggestions for improvement – including suggestions for preventive actions – are also documented. The procedures for managing these suggestions are described in general terms in a general instruction [17] and more specifically in the internal procedure of the program.

12. Conditions and obligations

The general conditions for participation and the obligations of the participating laboratories and the organiser are stated on our website: www2.slv.se/absint.

Contracts with special conditions and obligations can be established between the organiser and an individual laboratory when necessary.

12.1. Extract of the general conditions of participation

12.1.1. Who can participate?

- Laboratories that perform analyses within the frames of the schemes and that are using relevant methods.
- Laboratories to which consignments will be available in time by use of ordinary mailing facilities or a carrier service, and that are able to report results and pay invoices in due time.
- Laboratories with internet access, and that are able to use the website of the program.

12.1.2. Which methods may be used?

- All methods adapted for the analytical parameters that are evalauted in the PT schemes. The methods should, preferentially, be used as routine methods.

12.1.3. Fee

- An invoice is sent for the PT round(s) the laboratory has participated in.
- The fee shall be paid within the time period stipulated (normally within 30 days after print-out of the invoice).

12.2. Other obligations of participating laboratories

- To visit the program website and actively register/unregister for participation in the PT rounds.
- To report results according to the written instructions.

12.3. Obligations of the Swedish Food Agency

- To keep the information on the website up-to-date, e.g. information regarding PT rounds, analyses, dates and prices.
- To publish the original and preliminary processed results on the website within the stated period of time (see 8.7.1.).
- To publish a final report as a pdf-document on the website within the stated period of time (see 8.7.2.).

12.4. Limited responsibility

- The provider has no liability regarding third party claims depending on a laboratory's participation and performance in any of the schemes run by the Swedish Food Agency.

13. Participation costs

Current prices for participation in the respective PT rounds are stated on the website. The Swedish Food Agency reserves the right to change the prices if necessary, in order to be able to continue the activity based on the set requirements.

The fee for participation is paid after invoice. The prices are stated in Swedish crowns (SEK) as well as USD ($) and Euro (€) and payment can be made in any of these currencies.

14. This protocol

This protocol will be revised when important modifications or organisational changes are made. Participants will be notified whenever a new edition is available.

15. References

1. EN ISO/IEC 17025:2017. General requirements for the competence of testing and calibration laboratories. *(National versions issued 2018)*
2. EN ISO/IEC 17043:2010. Conformity assessment – General requirements for proficiency testing (ISO/IEC 17043:2010, IDT)
3. Peterz, M. & Steneryd, A.C. 1993. Freeze-dried mixed cultures as reference samples in quantitative and qualitative microbiological examinations of food. J. Appl. Bacteriol. 74:143-148.
4. International Standard, ISO 8199:2018. Water quality – General requirements and guidance for microbiological examinations by culture.
5. Ordinances of Swedish Work Environment Authority (*Arbetsmiljöverkets författningssamling*), 2018. Risks of infection *(Smittrisker)*. AFS 2018:4. *In Swedish*.
6. Kallings, I. & Ljungdahl Ståhle, E. Swedish Institute for Infectious Disease Control (*Smittskyddsinstitutet*), 2002. Expert's report regarding proficiency testing and reference samples from the Swedish Food Agency. The Public Health Agency of Sweden (*Fokhälsomyndigheten*) Reg. nr. 527/2002-18. Swedish Food Agency Reg. nr. 2509/02. *English translation.*
7. Frändberg, E. & Anér, G. Swedish Food Agency (*Livsmedelsverket*), 2002. Judgement on microbiological material for controls. Swedish Food Agency Reg. nr. 2509/02. *In Swedish*.

8. Thompson, M. & Wood, R. 1993. The international harmonized protocol for the proficiency testing of (chemical) analytical laboratories. Pure & Appl. Chem., Vol. 65, No. 9 2123-2144. *(Also published 1993 in: J. AOAC International 76, 926-940)*
9. Thompson, M., Ellison, S. & Wood, R. 2006. The international harmonized protocol for the proficiency testing of analytical chemistry laboratories. Pure & Appl. Chem., Vol. 78, 145-196.
10. ISO 13528:2016 Statistical methods for use in proficiency testing by interlaboratory comparison.
11. BCR Information, Chemical analysis.1993. Statistical analysis of certification trials for microbiological reference material. Commission of the European Communities, EUR 15008 EN.
12. Elliott, J. M. 1977. Some methods for statistical analysis of samples of benthic invertebrates. Freshwater Biological Association, Scientific publication no. 25, 2nd edition, 4th impression 1993.
13. Mooijman, K. M., During, M. & Nagelkerke, N. J. D. 2003. MICROCRM: Preparation and control of batches of microbiological materials consisting of capsules. RIVM report 250935001/2003. RIVM, Bilthoven, Holland.
14. Peterz, M. 1992. Proficiency testing: external evaluation of analytical competence in Microbiological Food laboratories. SLV rapport nr 21, 1992, 15 p., Swedish Food Agency. *In Swedish*.
15. Grubbs, F. & Beck, G. 1972. Extension of sample sizes and percentage points for significance tests of outlying observations. Technometrics 14:847-854.
16. Kelly, P. 1990. Outlier detection in collaborative studies. J. Assoc. Off. Anal. Chem. 73:58.64.
17. Swedish Food Agency (*Livsmedelsverket*), The Quality system, Science Division. “Actions in relation to discrepancies, complaints and proposals for improvement”. Latest version in the quality manual on the intranet of the Swedish Food Agency. *In Swedish.*

Proficiency Testing

SWEDAC ACKREDITERING
1457
ISO/IEC 17043

Livsmedelsverket
Swedish Food Agency

PT micro
Since 1981