MINISTRY OF THE ENVIRONMENT,
CONSERVATION AND PARKS

Ontario

Technical Guidance

Protocol for Evaluations of RT-qPCR Performance Characteristics

January 2022

ISBN 978-1-4868-5481-3

The development of this document was supported by ONTARIO CLEAN WATER AGENCY
AGENCE ONTARIENNE DES EAUX

Protocol for Evaluations of RT-qPCR Performance Characteristics

Protocole d'évaluation des caractéristiques de rendement des PCR quantitatives en temps réel

Technical Guidance

Guide technique

January 2022

Janvier 2022

Acknowledgements

The *Protocol for Evaluations of RT-qPCR Performance Characteristics* ("Protocol") was prepared for the Ontario-led COVID-19 Wastewater Surveillance Initiative (WSI).

The development of this Protocol was initiated and led by Alex Ho Shing Chik, Ph.D. (Ontario Clean Water Agency [OCWA]), and it was supported by Ontario WSI and Canadian wastewater surveillance stakeholders.

This Protocol was developed by various branches of the Ministry of the Environment, Conservation and Parks (MECP) and the following contributors, stakeholders and reviewers.

- Ann-Marie Abbey, M.Sc., MECP
- Alex Ho Shing Chik, Ph.D., OCWA
- Cecily Flemming, M.Sc. , MECP
- Tim Fletcher, MECP
- Jane Ho, P.Eng., OCWA
- Moustapha Oke, Ph.D., MECP
- Vince Pileggi, Ph.D., P.Eng., MECP
- Melanie Raby, Ph.D., MECP
- Albert Simhon, Ph.D., MECP
- Janis Thomas, Ph.D., MECP
- Susan Weir, Ph.D., MECP

Contributors and Stakeholders

- Patrick D'Aoust, M.Sc., University of Ottawa
- Stephen Brown, Ph.D., Queen's University
- Tzu-Chiao Chao, Ph.D., University of Regina
- Jade Daigle, M.Sc., Public Health Agency of Canada – National Microbiology Laboratory
- Robert Delatolla, Ph.D., University of Ottawa
- Hadi Dhiyebi, M.Sc., University of Waterloo
- Michael Donaldson, Ph.D., Trent University
- Elizabeth Edwards, Ph.D., University of Toronto
- Monica B. Emelko, Ph.D., University of Waterloo
- Qiudi Geng, Ph.D., University of Windsor
- Kimberley Gilbride, Ph.D., Ryerson University
- Tyson Graber, Ph.D., Children's Hospital of Eastern Ontario
- Christopher Kyle, Ph.D., Trent University
- Chand Mangat, Ph.D., Public Health Agency of Canada – National Microbiology Laboratory
- Kevin McDermott, Public Health Ontario
- Mike McKay, Ph.D., University of Windsor
- Claire Oswald, Ph.D., Ryerson University
- Xiao-Li (Lilly) Pang, Ph.D., University of Alberta
- Hui Peng, Ph.D., University of Toronto
- Natacha Prostran, B. Sc., University of Ottawa
- Philip J. Schmidt, Ph.D., AStat, University of Waterloo

- Mark Servos, Ph.D., University of Waterloo
- Nivetha Srikanthan, M.Sc., University of Waterloo
- Amir Tehrani, Ph.D., Ryerson University
- Ivy Yang, Ph.D., University of Toronto
- Yuwei Xie, Ph.D., University of Saskatchewan

External Peer Reviewers

The authors are grateful to the following people for reviewing an earlier version of the *Protocol.*

- Warish Ahmed, Ph.D., CSIRO, Australia
- Aaron Bivins, Ph.D., University of Notre Dame, USA
- Andreas Farnleitner, Ph.D., Karl Landsteiner University of Health Sciences, Austria
- Hannah Greenwald, University of California, Berkeley, USA
- Susana Guix, Ph.D., University of Barcelona, Spain
- Charles Haas, Ph.D., Drexel University, USA
- Rose Kantor, Ph.D., University of California, Berkeley, USA
- Ana Maria de Roda Husman, Ph.D., RIVM, The Netherlands
- Jérôme Lok-Wah-Hoon, M.Sc., LIBMS, RIVM, The Netherlands
- Joris Sprokholt, Ph.D., RIVM, The Netherlands

Disclaimer

The Ontario Clean Water Agency and the Ministry of the Environment, Conservation and Parks (MECP) have prepared this guidance manual. The *Protocol* has been subjected to these organizations' peer and administrative review and has been approved for publication as an MECP document.

This *Protocol* is not a formal policy or regulatory requirement; MECP offers it as guidance for laboratories deploying analyses based on reverse transcription quantitative polymerase chain reaction-based (RT-qPCR) in environmental samples and for decision makers who need to judge the quality of RT-qPCR data. This Protocol has been prepared considering SARS-CoV-2 as a primary target of interest; however, the concepts and theory presented are broadly applicable to other targets to be evaluated in environmental samples and other complex matrices.

The mention of trade names or commercial products in this manual does not constitute endorsement or recommendation for use.

We consider this document a living document intended to reflect the latest knowledge on the subject, but due to the pace of change in the field, this document may not reflect the latest knowledge on any particular issue. We welcome questions, comments and recommendations related to this document and these can be forwarded to us by sending an email to wastewater.surveillance@ontario.ca. We ask that you include all relevant and related supporting information, with key reference documents cited, to assist us in addressing the issues raised expeditiously.

Foreword

The surveillance of wastewater for SARS-CoV-2 to monitor COVID-19 within populations is underway in jurisdictions globally. The presence and quantity of SARS-CoV-2 genetic material in wastewater offers the potential to monitor a broader portion of the population to identify community infections without relying solely on clinical samples. Testing wastewater can capture both symptomatic and asymptomatic people, helping to detect the presence and trends of COVID-19 infections in a community setting. These results can complement clinical and public health data by identifying 'hot spots' for the virus and in some instances can help to identify infections ahead of clinical data. This information can help to inform public health decisions on where and how to mobilize resources.

Starting in fall 2020, the Province of Ontario partnered with academic and research institutions, in cooperation with Public Health Units and municipalities, to create an integrated program that expanded wastewater sampling and analysis province-wide. The Wastewater Surveillance Initiative (WSI) includes testing at wastewater treatment plants across the province and also includes vulnerable populations such as First Nation communities, long-term care homes, shelters and correctional facilities.

While a standard method is not currently established for SARS-CoV-2 testing in wastewater, efforts have been made by the province and participating laboratories to establish many common practices and protocols. As well, all participating laboratories have completed at least one inter-laboratory study to compare results between participants and to allow for further enhancements to methodologies within and between laboratories. Without a standard method, establishing and performing adequate quality assurance/quality control (QA/QC) measures is critical for laboratories to ensure high quality data are produced to support public health initiatives in the province.

Executive Summary

The generation of reliable results from analytical methods to quantify SARS-CoV-2 RNA in wastewater is paramount to the Wastewater Surveillance Initiative (WSI) as it forms the basis for assuring Public Health officials that the information is useful and potentially actionable when used to complement other epidemiological metrics. Because there is no standardized method for analysis of SARS-CoV-2 RNA in wastewater, the demonstration of method performance through the implementation of stringent quality assurance/quality control (QA/QC) protocols is critical for assuring the reliability of the data. As wastewater surveillance efforts have matured from individual proof-of-concept research studies to coordinated efforts across multiple scales and jurisdictions globally, there is a growing need for an alignment of definitions of performance characteristics to allow for their meaningful evaluation and comparison. Although minimum reporting needs for qPCR-based methods have been outlined by *Minimum Information for Publication of Quantitative Real-Time PCR Experiments* (MIQE; Bustin et al., 2009) and other works, the broad range of pre-treatment steps that are often deployed to quantify qPCR targets in wastewater (and other environmental matrices) introduce additional considerations for data generation, analysis, and interpretation.

Inter-laboratory studies conducted in Canada and internationally focusing on RT-qPCR-based methods for the quantification of SARS-CoV-2 RNA in wastewater have particularly emphasized 1) the need for a robust standard curve developed using properly prepared and quantified materials, 2) the need for various positive and negative controls to confirm the validity of analytical results, and 3) the need for accurate and transparent portrayal of the analytical results along with key performance characteristics for data reporting, analysis and interpretation. These critical elements form the basis of the framework herein described in this *Protocol*, as illustrated in Figure 1 below.

In an effort to prevent avoidable bias attributable to the use of different standard materials, the material(s) used for development of standard curves must be rigorously quantified and verified; the use of a single reference material with certified quantities to benchmark other standards used for quantification of SARS-CoV-2 RNA is herein recommended. Verification of quantities can be done using an independent method prior to use of the standards for quantification; digital PCR (dPCR) offers one such possibility. Nucleic acid quantification assays (e.g., Invitrogen Qubit™ or Thermo-Fisher NanoDrop™) can also be used to characterize reference material. As the standard curve serves as the calibrator against which any and all signals present in samples are compared, the robustness, repeatability and reproducibility of the standard curve is of paramount importance and must be ascertained in accordance with stringent acceptance criteria.

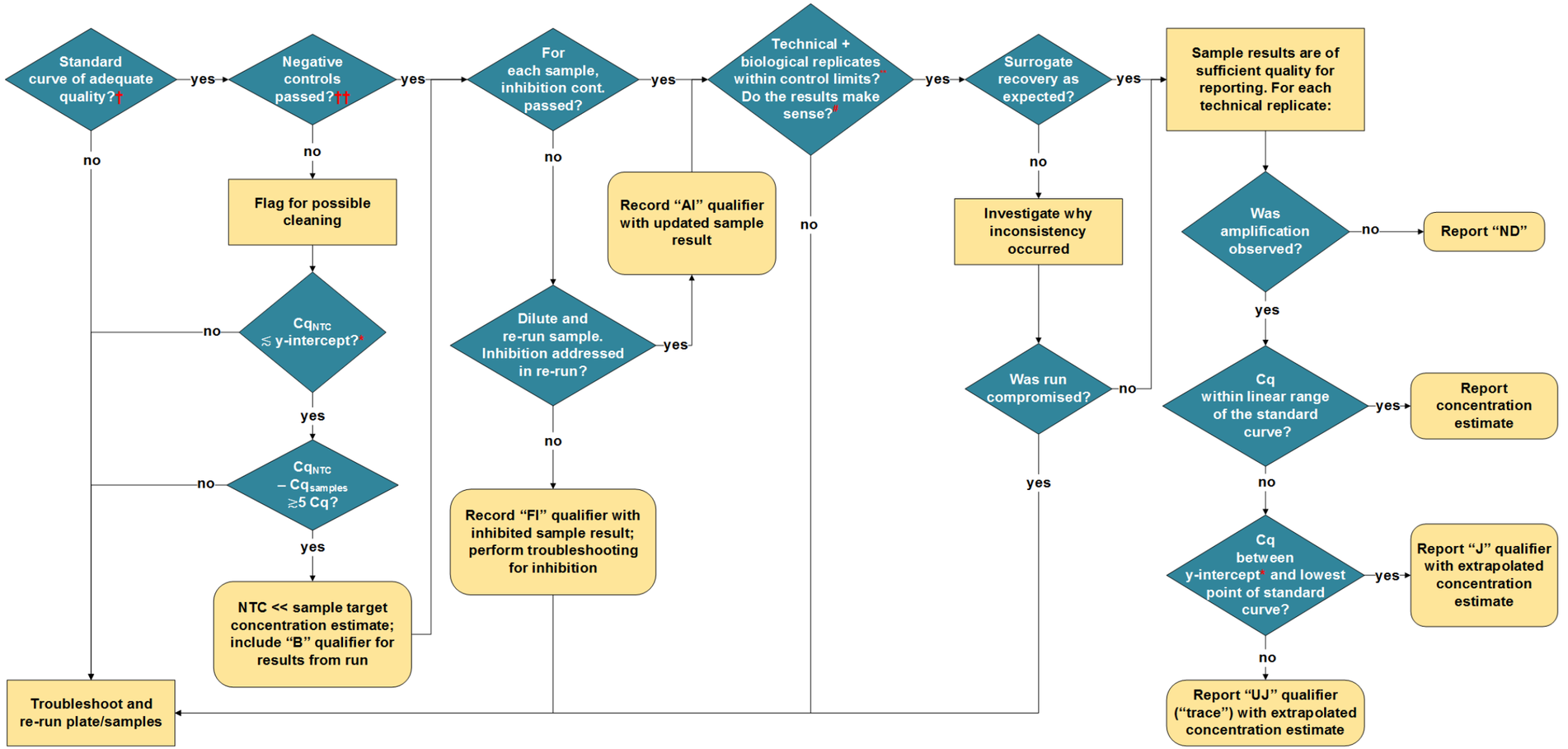
False-positive and false-negative results can undermine the validity of sample results. Therefore, appropriate positive (e.g., extraction controls, internal amplification controls) and negative (e.g., whole process blanks, sample blanks) controls should be deployed. Wastewater matrices pose a challenge as they can contain various inhibitory substances that can impede or completely prevent PCR reactions from occurring, resulting in a biased estimate of qPCR target concentrations. Mutations in the binding sequences may also impact assay sensitivity due to the incompatibility of the primer and probe sets used and the target of interest. The concordance between two or more target regions (e.g., N1, N2, E genes of SARS-CoV-2) can mitigate this risk and the tracking of more than one target region is therefore recommended. When an issue is

noted, appropriate troubleshooting steps must be taken to rectify the problem and be clearly documented.

Target concentration estimates determined using RT-qPCR are inherently variable and subject to uncertainty. However, expectations related to assay sensitivity, specificity, precision, and recovery, as well as the behaviour of the method considering the wastewater matrix processed, must be established and documented to avoid further bias to the analytical result. The laboratory analyst must develop and maintain rigorous documentation of control limits and other QA/QC criteria; this traceability will help with troubleshooting the various steps involved throughout sample processing (e.g., random sampling error, aliquoting, pipetting and method losses). When wastewater surveillance is deployed to track temporal trends at a specific location, antecedent results can also provide an additional check of whether the latest analytical results make sense.

Upon ascertaining that sample results obtained are of sufficient quality, the analytical result must be portrayed to reflect potential limitations of the result and assay. Sample results must be reported with pertinent qualifiers that can consistently and transparently convey such limitations. Important information such as the initial sample process volume, the effective sample volume represented by the assay, and the sample limit of detection/quantification (SLOD/SLOQ, determined as outlined in this *Protocol*) must also be reported with the analytical result such that the development and application of appropriate statistical models for data analysis and interpretation can be facilitated.

At the time that this *Protocol* was prepared, method development and optimization for the quantification of SARS-CoV-2 in wastewater remains an active, ongoing area of research. Accordingly, substantial scientific uncertainty still persists for several aspects of the methodologies deployed for wastewater surveillance. In particular, active, ongoing research is underway with respect to 1) the selection, validation, and quantification of various fecal indicators to normalize the SARS-CoV-2 RNA signal for wastewater strength, and 2) the selection, validation, administration, and interpretation of various surrogates deployed to evaluate the recovery of SARS-CoV-2 in wastewater using various methodologies. Therefore, this *Protocol* may need to be updated in light of emerging knowledge and method advancements and has been prepared in a scientifically-informed and comprehensible structure that would better facilitate the “fit-for-purpose” adaptation of such emerging knowledge.



Abbreviations:

Cq	Quantification cycle
NTC	No template control

List of qualifiers (multiple qualifiers can pertain to each sample):

B	Background contamination observed
FI	Failed inhibition control
AI	Addressed inhibition
ND	Non-detect
J	Concentration estimate extrapolated based on extending experiment-specific standard curve to the y-intercept
UJ	"Trace" amplification of target; concentration estimate extrapolated based on extending experiment-specific standard curve beyond the y-intercept

All results must be reported with the concentration factor (CF), effective sample volume (ESV), and the original sample volume processed.
If surrogate recovery was evaluated, the estimated surrogate recovery (%) should be reported with the recovery surrogate used.

Notes:

† Refer to section on standard curves and Figure 3.3 in this Protocol

†† Refer to section on negative controls in this Protocol

* The experiment-specific standard curve shall be used

** Control limits as established through control charting for positive controls; repeatability across technical replicates should also be verified

When the purpose of monitoring is to track temporal/spatial trends, antecedent/previous results can also provide an additional check of whether the latest analytical results make sense

Figure 1. Framework for evaluations of RT-qPCR performance characteristics: Implications for data handling, reporting and interpretation.

Table of Contents

List of Tables

List of Figures

1. Purpose of this Protocol

Evaluation and documentation of RT-qPCR performance characteristics is a critical component of quality assurance/quality control (QA/QC) because it provides information related to the assay and method robustness and reproducibility. These performance characteristics include, but are not limited to, the determination of thresholds related to precision and sensitivity. However, despite the recognition of their critical role, there remains a lack of alignment for key definitions that will allow for streamlined and meaningful evaluation of these characteristics. Minimum reporting needs have been outlined by *Minimum Information for Publication of Quantitative Real-Time PCR Experiments* (MIQE; Bustin et al., 2009); however, the quantitation of qPCR targets from wastewater and other environmental matrices can further be complicated by the broad range of pre-treatment steps deployed and matrix effects. The *Environmental Microbiology Minimum Information* Guidelines (EMMI; Borchardt et al., 2021) address some of these additional considerations for environmental applications of qPCR and dPCR. However, inconsistent definitions proposed for performance characteristics and concepts (such as assay "limit of detection" [LOD] and assay "limit of quantification" [LOQ]) in these and other reference works also often lack clear operational guidance (e.g., specification of the type of regression used to interpolate the assay LOD in ISO 20395:2019) or consideration of their practical utility (e.g., when a consistently high level of precision is achieved and interpolation of an LOQ threshold is either impossible or not meaningful). These considerations could confound the interpretation of analytical results generated using these methods if not streamlined. Therefore, a need to go beyond minimum reporting requirements to adequately and transparently portray the quality of the data reported is underscored.

The purpose of this document ("*Protocol for Evaluations of RT-qPCR Performance Characteristics*", herein referred to as "*Protocol*") is to provide some operational definitions and delineate a streamlined approach—wherever reasonable and feasible—for determining these performance characteristics. Key distinctions of performance characteristics evaluated at the level of the assay (considering the RT-qPCR reaction only) or the sample (considering sample processing steps leading to and including the RT-qPCR) are also highlighted herein. Accordingly, the performance characteristics described in this *Protocol* pertain exclusively to the analytical method applied. While appropriate sampling program design and sample collection procedures also play critical roles in the generation and interpretation of the analytical results, these considerations are outside the scope captured by this *Protocol*. It must be emphasized that the streamlining of QA/QC protocols for RT-qPCR methods to estimate SARS-CoV-2 concentrations in wastewater does not imply that the methods themselves are streamlined; the reconciliation of results emanating from different methods remains a key challenge and an ongoing area of research. Finally, this Protocol has been prepared focusing on SARS-CoV-2 as a primary target of interest; however, the concepts and theory presented stem from the analysis of other targets evaluated in environmental samples and other complex matrices, and are therefore broadly applicable.

This Protocol is developed based on the most current available knowledge and is intended for an audience that includes technical field and laboratory personnel with prior experience in performing RT-qPCR analyses. This Protocol is to be used with professional judgement and experience related to wastewater field sampling campaigns and laboratory wastewater analysis. If there is any doubt in the use of this Protocol, please seek professional advice.

2. Definitions

Definitions for common terminology related to RT-qPCR performance characteristics are provided below:

Performance characteristics

Accuracy: refers to the difference between experimentally measured and actual concentrations, presented as fold changes or copy number estimates (Bustin et al., 2009).

Analytical sensitivity: probability of detection when the target is actually present in the source; proportion of samples with target present that are deemed positive. Analytical sensitivity can be described as a "limit of detection" (LOD) at several levels when applied to RT-qPCR-based methods:

> ***Assay limit of detection (ALOD):*** the minimum level of a target consistently detectable (e.g., with 95% probability) using the assay considering only the amplification and quantification steps of RT-qPCR.
>
> ***Sample limit of detection (SLOD):*** the minimum level of a target within a sample that would be consistently detectable considering all sample processing steps, from sample concentration leading up to and including amplification and quantification steps of RT-qPCR. In this Protocol, the SLOD is the concentration factor-adjusted ALOD, which provides a theoretical estimate of the minimum concentration in the original sample that would yield consistent detection of the target, assuming no method losses.
>
> ***Method limit of detection (MLOD):*** the minimum level of a target necessary for its consistent detection considering all sample processing steps (as in the definition of SLOD), but considering biases attributable to method losses. The MLOD is experimentally determined using a rigorous and iterative testing approach and it is sensitive to the recovery surrogate used and the sample matrix encountered.

Specificity: the degree to which an assay is specific for a particular target; ability of a measurement procedure to measure solely the target analyte (International Organization for Standardization, 2019).

Precision: The closeness of repeated measurements to each other; intra-assay ***repeatability*** refers to the variance of the assay with the same samples repeatedly analyzed in the same assay, while inter-assay ***reproducibility*** refers to the variation in results between runs or different laboratories and is typically expressed as the standard deviation or coefficient of variation of copy numbers or concentrations (Bustin et al., 2009). The concept of "limit of quantification" (LOQ) can be used as a statistical measure of precision at various levels (i.e., assay, sample, method) analogous to those described for analytical sensitivity.

Recovery: the proportion of a specific target within an analytical portion of a test material (e.g., wastewater sample) that was successfully extracted and detected/represented by the final measurement.

Sample processing characteristics

Sample volume: the volume of sample input into the primary concentration method.

Concentration Factor (CF): the degree to which the concentration observed in the final qPCR assay volume has been magnified compared to concentration of the analyte in the original sample (Example in Box 2.1);

$$CF = \frac{V_{sample}}{V_{after\ concentration}} \times \frac{V_{concentrate\ for\ RNA\ extraction}}{V_{after\ RNA\ extraction}} \times \frac{1}{DF} \qquad \textit{(Equation 1)}$$

Where V_{sample} is the sample volume input into the primary concentration method [mL]

$V_{after\ concentration}$ is the sample volume after concentration [mL]

$V_{concentrate\ for\ RNA\ extraction}$ is the volume of the concentrate used for RNA extraction [mL]

$V_{after\ RNA\ extraction}$ is the total volume of RNA extracted [mL]

DF is the dilution factor of the template used in the RT-qPCR reaction (i.e., [diluent volume+RNA extract volume]/RNA extract volume; Pecson et al., 2021).

Where the solids fraction was analyzed, the proportion of the solids assayed should be used in place of the proportion of the volumetric concentrate assayed. Any dilutions of the RNA extract assayed to address inhibition should be reflected in the determination of the dilution factor.

Effective Sample Size/Volume (ESV): the amount of the original sample size/volume that was actually analyzed in a qPCR reaction (Example in Box 2.1).

Standard curve

Standard curve (also known as a calibration curve)***:*** A sample of known concentration units (e.g., pg/μL, copies/reaction, dilution factor, number of cells, or a relative dilution factor) is serially diluted through a controlled series. The observations or measurements, in this case Cq values of these standards, are plotted against the logarithm of their concentration and used to construct a standard curve. The standard curve is used to estimate analyte concentration of the unknown test samples from the observed Cq. (Nolan et al., 2013).

Coefficient of Variation (CV): A statistical measure of the relative dispersion of data points; used to express the repeatability and reproducibility of an assay. The coefficient of variation (CV_{ln}) must be calculated based on the concentration estimates rather than the Cq directly.

Dynamic range (also known as Linear range): The minimum and maximum range of values over which a given assay is linear (e.g., the highest to the lowest quantifiable copy number established by means of a calibration curve; Bustin et al., 2009).

PCR amplification efficiency: Amplification efficiency is determined from the slope of the log-linear portion of the calibration curve. Specifically, PCR efficiency = $10^{-1/slope} - 1$ when the logarithm (base 10) of the initial template concentration (the independent variable) is plotted on the x axis and C_q (the dependent variable) is plotted on the y axis. The theoretical maximum of 1.00 (or 100%) indicates that the amount of product doubles with each PCR thermal cycle. Ideally, the confidence intervals or standard errors of the mean of estimated PCR efficiencies are reported from replicated calibration curves (Bustin et al., 2009).

Threshold/quantification cycle (Ct/Cq): The number of amplification cycles using quantitative reverse transcription Polymerase Chain Reaction (RT-qPCR) required for the signal associated with a PCR product (i.e., the target/amplicon) to be detected above a baseline signal that would be present in the assay regardless of whether the target is present. Threshold cycle and quantification cycle have been interchangeably used (as have Ct and Cq denotations) in the literature; however, to avoid confusion with the automatic/manual instrument thresholds, the streamlined use of Cq rather than Ct has been proposed (Nolan et al., 2013).

QA/QC Controls

Technical replicates: repeated measurements of the same sample in the same assay that represent independent measures of the random noise associated with protocols or equipment. In this *Protocol*, technical replicates refer to replicate aliquots of the template obtained *from the same RNA extract* obtained from a sample that is run in the same experiment.

Biological replicates: the analysis of distinct samples collected representing an identical time point or condition. In this *Protocol,* biological replicates refer to aliquots of the same sample that are introduced at the initial processing step (e.g., 50 mL aliquots from a well-mixed 1.0-L sample; each 50 mL aliquot is subjected to the entire analytical procedure).

Negative control: the analysis of each primer set to verify that no contaminating nucleic acid has been introduced into reagents or into samples during sample processing (US EPA, 2004). Various other negative controls can be deployed at various steps of sample processing to troubleshoot contamination issues (Borchardt et al., 2021)

> ***No Template Control (NTC):*** qPCRs include all PCR reagents with the exception of the template (i.e., pipetting 5 µL of PCR-grade water and 20 µL of respective master mix into a well). This is a standard negative control used to identify set-up contamination and primer-dimer product amplification (Nolan et al., 2013).
>
> ***Whole process blank:*** A negative control initiated at the primary concentration step and carried through the entire analytical workflow to verify that no contamination has been introduced throughout the entire sample processing (US EPA, 2004).

Positive control: controls prepared and analyzed to verify that the method is capable of adequately recovering and amplifying the target (US EPA, 2004). Various other positive controls can be deployed at various steps of sample processing to evaluate recovery (Borchardt et al., 2021)

> ***Inhibition control:*** Inhibitory substances may be present that impede or prevent PCR from running efficiently or effectively, ultimately resulting in delayed Cq quantification (higher Cq) for the actual target of the analysis. Inhibition effects can be monitored by comparing the number of cycles required for detecting a target in a spiked sample matrix compared to that of a distilled water control spiked at the same concentration; others verify an expected linear decrease in signal after diluting samples.
>
> ***Internal Amplification Control (IAC):*** a control performed to verify successful amplification of targets, to confirm that negative results are not due to unsuccessful amplification (US EPA, 2004); this can be a target secondary to the target of interest (e.g., Pepper Mild Mottle Virus [PMMoV], crAssphage) to verify the presence of RNA in the sample and/or presence of inhibitors.

PCR positive control: a control performed to verify that the PCR master mix and reagents were prepared correctly to produce amplification of the target nucleic acid (US EPA, 2004); a synthetic or otherwise verified RNA template is used to verify that the RT-PCR assay was performed correctly (e.g., SARS-CoV-2 RNA standard).

Whole process control/Matrix spike/Recovery surrogate: A substance that usually shares characteristics of the target of interest and therefore, when assayed, is assumed to exhibit similar behaviour in response to the analytical procedure. It is used to evaluate the effects and responses to selected processing treatments and the effect of the matrix on method recovery; can be used to confirm inhibition effects. Whole process controls are usually spiked at the beginning of the sample processing procedure (prior to sample concentration) to facilitate the quantitative estimation of analytical recovery.

Box 2.1 Relationship between sample volume, concentration factor, effective sample volume

A protocol concentrates an initial sample volume of 40 mL to a volume of 0.1 mL. The full volume of the concentrate was introduced to an elution column; the RNA was eluted from the column to an elution volume of 0.1 mL. Half of the eluate was diluted (1 part eluate: 4 parts diluent) was necessary to address inhibition. 5 µL of the diluted eluate is used in each qPCR reaction with 15 µL of master mix (final qPCR reaction volume = 20 µL). A sample was processed accordingly and after quantification by standard curve, the sample exhibited a Cq that corresponded with a signal consistent with 15 gene copies/reaction.

Given:

- Sample volume: 40 mL
- Concentrate volume: 0.1 mL
- Eluate volume: 0.1 mL; 50 µL of eluate used for dilution
- Dilution factor to address inhibition: (200+50)/50 = 5
- Volume of diluted eluate used for qPCR reaction: 5 µL
- qPCR final reaction volume: 20 µL

$$CF = \overbrace{\frac{V_{sample}}{V_{after\ concentration}}}^{\text{concentration step}} \times \overbrace{\frac{V_{concentrate\ for\ RNA\ extraction}}{V_{after\ RNA\ extraction}}}^{\text{elution step}} \times \overbrace{\frac{1}{DF}}^{\text{dilution step}} \qquad \textit{(Equation 1)}$$

$$= \frac{40\ \text{mL}}{0.1\ \text{mL}} \times \frac{0.1\ \text{mL}}{0.1\ \text{mL}} \times \frac{1}{5} = 80$$

The concentration factor is 80, meaning that the reaction concentration estimate (15 gene copies in 5 µL of template) is 80 times that of the concentration of the original sample. Therefore, the concentration estimate of the original sample is estimated to be 37.5 gene copies/mL.

N.B. (1): Some laboratories may be accustomed to using the concentration estimate on a volumetric basis in the reaction well (15 gene copies in 0.020 mL = 750 gc/mL). This value will only be 20 times relative to the concentration of the original sample [i.e., CF (80) × ratio of the diluted eluate volume relative to the qPCR final reaction volume (5/20)].

The effective sample volume (ESV) is the amount of the original processed volume that has been proportionally represented by the final volume assayed. Each term represents the proportion of volume derived from the preceding step:

$$ESV = V_{sample} \times \overbrace{\frac{V_{concentrate\ for\ RNA\ extraction}}{V_{after\ concentration}}}^{\text{eluate}} \times \overbrace{\frac{V_{eluate\ used\ in\ dilution}}{V_{after\ RNA\ extraction}}}^{\text{diluted eluate}} \times \overbrace{\frac{V_{diluted\ eluate\ assayed}}{V_{diluted\ eluate}}}^{\text{assay}} \qquad \textit{(Equation 2)}$$

$$= 40\ \text{mL} \times \frac{0.1\ \text{mL}}{0.1\ \text{mL}} \times \frac{0.05\ \text{mL}}{0.1\ \text{mL}} \times \frac{5\ \mu\text{L}}{250\ \mu\text{L}} = 0.4\ \text{mL}$$

The signal from the well (15 gene copies/reaction) can be divided by the ESV (0.4 mL) to determine the concentration estimate of the original sample (15 gene copies/0.4 mL = 37.5 gene copies/mL).

The determination of CF and ESV is particularly important for inter-laboratory method comparisons, as well as interpretation of analytical results when target concentrations are expected to be low and likely subject to random sampling error.

N.B.(2): the $V_{after\ concentration}$ and $V_{concentrate\ for\ RNA\ extraction}$ can be replaced with the M_{pellet} (wet mass of the pellet) and the $M_{pellet\ for\ RNA\ extraction}$ (wet mass of the portion used for RNA extraction), respectively, for methods that rely on the generation and subsampling of a pellet.

N.B.(3): a mass balance (law of conservation of mass) performed on solids content parameters (e.g., total suspended solids) may be a necessary consideration for the determination of the CF and ESV if a step involving the pre-concentration/sampling of the solids in a wastewater sample was performed (e.g., targeted analysis of settled solids from a wastewater sample).

3. Standard curves for RT-qPCR

Choice of standard material for the quantification of SARS-CoV-2

In an effort to prevent avoidable bias attributable to the use of various standard materials, the material(s) used for development of standard curves must be rigorously quantified and verified; the use of a single reference material with certified quantities to benchmark other standards used for quantification of SARS-CoV-2 RNA is herein recommended. While costs, supply chain limitations, and laboratory protocols might impose challenges for the routine use of a single standard material with certified quantities, it is prudent that a single such material common to WSI participants be used at least occasionally to facilitate methods comparisons as well as provide an additional indicator of quality control.

- Based on a review of current literature (Figures 3.1 and 3.2), synthetic RNA standards have emerged as the most commonly used type of standard used amongst labs internationally that partake in wastewater surveillance efforts for SARS-CoV-2 (Bivins et al., 2021). Other standards exist and are in use, including linearized plasmid DNA standards and DNA standards (e.g., IDT gBlocks™), as well as gamma irradiated SARS-CoV-2—each with their advantages and disadvantages along with unique handling and preparation requirements. When handled, stored, and prepared properly in accordance with the manufacturer's instructions, they have been noted to perform comparably against RNA standards.

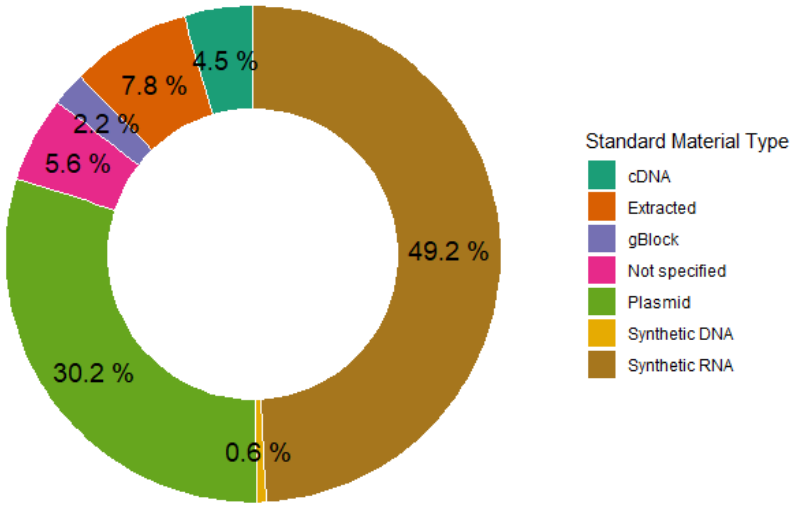


Figure 3.1. Review of standard material types reported to be used in SARS-CoV-2 quantification, amongst 169 published studies. A total of 179 standards were documented (some studies utilized more than a single standard).

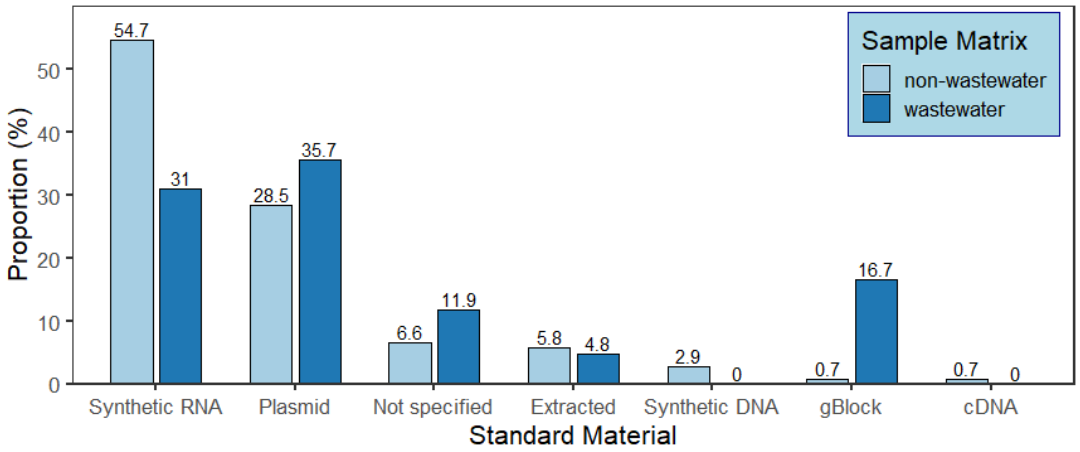


Figure 3.2. Proportion of standard types used amongst 169 published studies reviewed, by the type of sample matrix examined

- However, a key benefit of the use of RNA standards is that their use implicitly captures the efficiency of reverse transcription associated with one-step RT-qPCR (when both reverse transcription and amplification are performed in the same vessel), to which the RNA of SARS-CoV-2 is also subject. Moreover, the use of RNA standards with certified copies of target material (obtained via a supplier's certification documentation or by empirical analysis at the participating laboratory) is essential as suppliers may not provide certification on a per-lot basis.

- Recognizing that our collective understanding of standard materials and methods has evolved and will continue to evolve, we encourage all laboratories participating in the WSI to routinely use a certified reference RNA template (e.g., EDX RNA standard or equivalent for which target RNA copy numbers have been quantified and certified for each production lot by the vendor; accompanied by a Certificate of Analysis). These standard materials (whether commercial or home-made DNA/RNA standard materials) should be quantitated using digital PCR (dPCR, e.g., ddPCR™) or benchmarked against a certified reference template RT-qPCR standard curve on a monthly basis, at a minimum.

- Certified reference RNA template (herein also referred to as a "certified RNA standard") should also be verified periodically to verify that reported concentrations are accurate. This may be achieved through the use of an independent verification of the concentration estimate using dPCR. It is recommended that a new certified RNA standard used at the lab be verified three times (to obtain a mean and standard deviation) and on a periodic basis thereafter (e.g., every other batch). Nucleic acid quantification assays (e.g., Invitrogen Qubit™ or Thermo-Fisher NanoDrop™) can also be used for IDT gBlock™ templates; however, this verification will be done on a mass basis rather than a most probable number concentration estimate of gene copies/mL basis. Approaches to verify the integrity of standards used are described in Annex B of ISO 20395:2019 (International Organization for Standardization, 2019).

- Laboratories benchmarking against a certified RNA standard on a monthly basis as described above that do not have access to dPCR can verify the accuracy of the certified

RNA standard during split sample testing conducted every other month with other labs that do have access to dPCR.

- This practice will be reviewed periodically to ensure that it continues to be consistent with the goals and intents of the best recommended practices.

Minimum requirements for standard curves

Labs participating in WSI were asked to conform with the following requirements for RT-qPCR standard curves: A minimum 5-point standard curve with at least duplicate analysis (triplicate preferable) for each point for every RT-qPCR experiment; or, standard curves from previous runs may be used provided the curve was created with a minimum of 6-points with 7 replicates and two of these points added with each qPCR run in duplicate as positive controls.

- MIQE guidelines stipulate that the dynamic range (linear portion) of the standard curve must include the interval for targets being quantified. It further recommended that the dynamic range (linear portion) of the standard curve span at least 3 orders of magnitude but ideally should extend to 5 or 6 $\log_{10}$ concentrations (Bustin et al., 2009).

- Laboratories should not proceed with a long-term standard curve with only positive controls every run (i.e., "simplified" standard curve) until characteristics of an "acceptable" standard curve has been developed over multiple runs (e.g., ≳30 standard curves)

- Standard curve slopes (and associated PCR efficiencies), as well as the coefficients of determination (r^2 values) should be tracked and reported. Ideally, confidence intervals should be estimated through the entire dynamic range (Bustin et al., 2009). Generally accepted PCR efficiencies are in the range of 90 and 110% (slope ranging from -3.1 to -3.6); and an r^2 value of at least 0.98 (Broeders et al., 2014; Svec et al., 2015). Information about the baseline threshold (manual or automatically set) and the y-intercept for each run should also be documented.

- A framework for determining when errors/outliers can be excluded, a standard curve needs to be re-run, and/or the entire assay must be re-run is presented in Figure 3.3. Please see the sections on intra-assay (repeatability) and inter-assay (reproducibility) precision below for further guidance. Interim thresholds for operator errors/exclusion of outliers have been proposed (Figure 3.3). The repeatability threshold (SD<0.5 Cq) should be applied at dilutions greater than ~10 gene copies/reaction. More stringent laboratory-specific criteria may be applied; the practicality and statistical implications of these interim thresholds will be reviewed to ensure an appropriate false rejection rate is achieved.

Qualitative vs. quantitative applications of RT-qPCR

- Circumstances may arise for the adaptation of RT-qPCR methods for qualitative applications (such as for the detection of SARS-CoV-2 from passive sampling devices that do not lend for quantitative interpretations). In such cases, performance criteria (e.g., acceptable PCR efficiencies) may differ to reflect the ultimate intended use and interpretations of the data (Broeders et al., 2014). For example, a broader range of PCR efficiencies has been deemed acceptable (80-120%) for qualitative purposes (Broeders et al., 2014). However, other performance characteristics must still be ascertained for

qualitative uses of RT-qPCR (e.g., determination of assay sensitivity). In some cases, a standard curve may not need to be generated; clear amplification of the target within the operational range of the instrument with appropriate negative and positive controls may suffice.

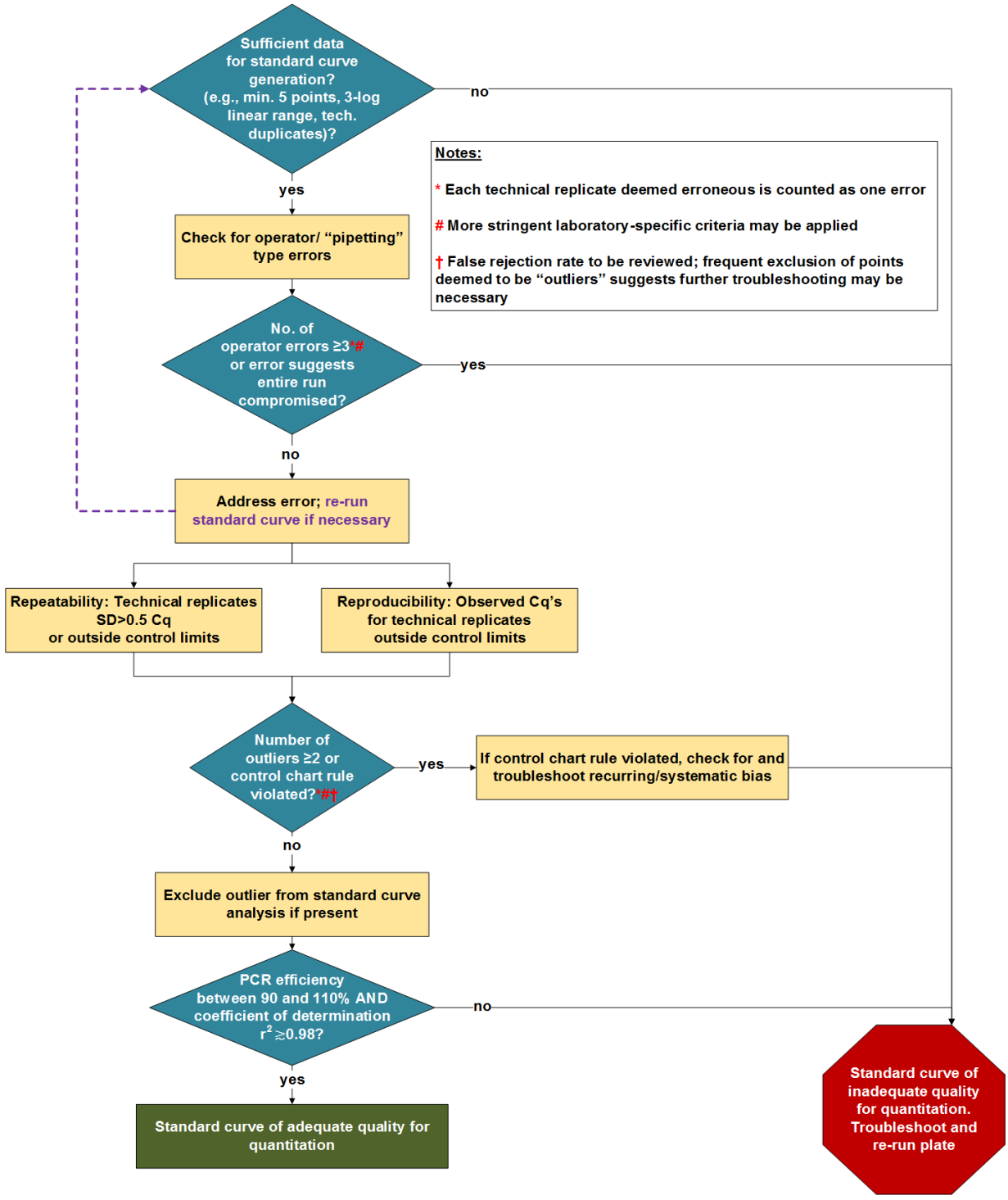


Figure 3.3. Framework for determining acceptability of standard curve for quantification

4. Analytical sensitivity

Analytical sensitivity refers to the ability to detect an RNA/DNA target when it is actually present in an assay; more specifically, for RT-qPCR assays, it relates to the minimum number of copies in a sample that can be consistently detected (i.e., true positive). This characteristic can be determined for the RT-qPCR assay using standard materials (assay sensitivity) **OR** based on targets in real wastewater matrices considering method losses (method sensitivity).

Mutations in the binding sequences may also impact assay sensitivity due to the incompatibility of the primer and probe sets used and the target(s) of interest. The concordance between two or more target regions (e.g., N1, N2, E genes of SARS-CoV-2) can mitigate this risk and the tracking of more than one target region is therefore recommended.

Assay sensitivity

- Although many definitions exist, the “assay limit of detection” (ALOD) can be defined as the lowest copy number of the target that yields a detectable PCR amplification product in all technical replicates tested, or in the majority of technical replicates (e.g., at least 95%) (Burns & Valdivia, 2008; Bustin et al., 2009; Nutz et al., 2011)
- Assay sensitivity for qPCR-based methods is subject to the limitations of the lowest count (1) of discrete targets present within a reaction volume.
- In absence of an independent method to confirm the number of targets present within the reaction volume, the number of targets present within a reaction volume will be assumed to be Poisson distributed with a mean based on the intended average concentration (Table 4.1):
    - At an intended average concentration of 5 targets (gene copies)/reaction, the probability that a reaction will contain 1 or more gene copies is 99.33%. There is a low probability (<1%) that a reaction volume prepared at this intended average concentration contained no gene copies (0), which will result in a non-detect as there are no templates to amplify (assuming there are no false-positives).
    - At an intended average concentration of 3 gene copies/reaction, the probability that a reaction will contain 1 or more gene copies is 95.02%. There is approximately 5% probability that a reaction volume prepared at this intended average concentration contained no gene copies (0), which will result in a non-detect as there are no templates to amplify.

Table 4.1. The influence of random sampling error on the number of targets that are present within a single reaction based on the intended average concentration.

		Probability(Number of targets in a single reaction)	
		P(0)	P(>1)
mean # of targets per reaction (gc/reaction)	0	1	0
	0.001	0.9990	0.0010
	1	0.3679	0.6322
	2	0.1353	0.8647
	3	0.0498	0.9502
	4	0.0183	0.9817
	5	0.0067	0.9933
	10	4.54×10^{-5}	0.99995
	49.5	3.18×10^{-22}	1
	100	3.72×10^{-44}	1
	1000	0	1

- The intended average concentration of 3 gene copies/reaction, considering only random sampling (Poisson) error, leads to at least one gene target in a reaction with 95% probability (numerically consistent with the ALOD definition above) and therefore has been widely regarded as the "theoretical LOD" (Bustin et al., 2009). However, this does not imply that targets, when present below this threshold, cannot be detected reliably. The random sampling error (also referred to as "subsampling error" (Taylor et al., 2019)) at such low target concentrations precludes meaningful/analogous use of the chemical definition of LOD as a descriptor of qPCR assay sensitivity. Accordingly, successful amplifications of targets below this threshold are valid results and should not be interpreted as though the analyte is absent; however, the probability of their repeated detection is expected to be less than 95% (Kralik & Ricchi, 2017).

- Increasing the number of technical replicates, the effective sample size captured by more efficient sample concentration, or the template volume used in the PCR reaction may be necessary to estimate the concentrations of samples with low levels of target (Taylor et al., 2019).

- To empirically determine an $ALOD_{95\%}$—the target mean concentration in a reaction well that would yield amplification in 95% of reactions— a *minimum* of six (6) replicates has been recommended (Kralik et al., 2011; Slana et al., 2008); however, a higher level of confidence can be achieved by performing more replicates (10 or 15 and more; Ricchi et al., 2016) using six (6) dilutions of the

standard material. At least one of the six dilutions should be set at a target concentration that will always yield amplification (i.e., no "drop-outs").

- The use of various probit/logistic regression approaches for determining an empirical ALOD have been proposed (Burns & Valdivia, 2008; Pavšič et al., 2015; World Organisation for Animal Health (OIE), 2014).
- Curve-fitting approaches that draw upon principles from dose-response models have also been proposed (Verbyla et al., 2016). Verbyla et al. (2016) applied an exponential model to specify the probability of amplification, which imposes more limitations on the shape of the fitted curve than probit/logistic regression. Until the assumptions implicit to this approach can be verified, we recommend using logistic regression in the interim.
- TAKE HOME MESSAGE FOR LABS: Most contemporary RT-qPCR assays have been suggested to be capable of consistently yielding detects (in 95% of reactions) at a target concentration of approximately 5 gene copies/reaction. A significant departure of the proportion of detects/non-detects for a given dilution of the standard material from the theoretical maximum (Box 4.1) would suggest that there is likely a source of variation in addition to subsampling error.
- To ensure the consistency and comparability of $ALOD_{95\%}$ determined within the WSI, it is proposed that all labs perform an experiment ("Repeatability Experiment") with a minimum of six (6) two-fold dilutions with the highest target concentration set at approximately ~30 gene copies/reaction (Box 4.1), similar to the approach suggested in ISO 20395:2019. When evaluating the $ALOD_{95\%}$ for a new method, a more "conservative" dilution series (i.e., higher starting target standard concentration $\gtrsim$30 gene copies/reaction) may be used as a starting point such that drop-outs only occur in 3 of the 6 dilutions. This might be desirable because this experiment may also inform the estimation of the $ALOQ_r$ (please refer to section on *Statistical Measures of Assay Precision)*. Fifteen (15) technical replicates are recommended for each dilution. Logistic regression will be used to empirically determine the $ALOD_{95\%}$. The experiment should be performed once every 6 months for each assay to confirm the consistency of the $ALOD_{95\%}$, at a minimum. Ideally, this experiment should also be run after calibration of an RT-qPCR instrument, a change of standard material and/or qPCR reagents. This value will need to be readjusted to the original processed sample volume using the concentration factor and reported as the $SLOD_{95\%}$. **The practicality and statistical implications of the $SLOD_{95\%}$ will need to be reviewed. A companion spreadsheet for the *Protocol* has been developed to support the determination of the $ALOD_{95\%}$ (and ALOQ) as described above.**
- Because of the random sampling error associated with discrete qPCR targets, the $ALOD_{95\%}$ determined is a relevant but incomplete indicator of assay sensitivity.

Box 4.1 Use of logistic regression to determine empirical $ALOD_{95\%}$

A lab performed a *Repeatability Experiment* with 15 technical replicates of standards prepared at 6 two-fold dilutions to evaluate the $ALOD_{95\%}$ empirically (Table 4.2).

Table 4.2. Example data collected from an experiment

Nominal standard concentration per reaction	0.9375	1.875	3.75	7.5	15	30
Log-10 gene copies per reaction	-0.028	0.273	0.574	0.875	1.176	1.477
Successful amplifications observed	3	5	14	15	15	15
Number of technical replicates performed	15	15	15	15	15	15

Table 4.3. Summary of logistic regression model (binomial regression with a logistic link function)

Coefficient	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.900	0.6144	-3.093	0.002
Log(Standard concentration per reaction)	6.843	1.608	4.257	2.07e-05

Null deviance: 60.8488 on 5 degrees of freedom
Residual deviance: 3.5629 on 4 degrees of freedom
AIC: 15.347

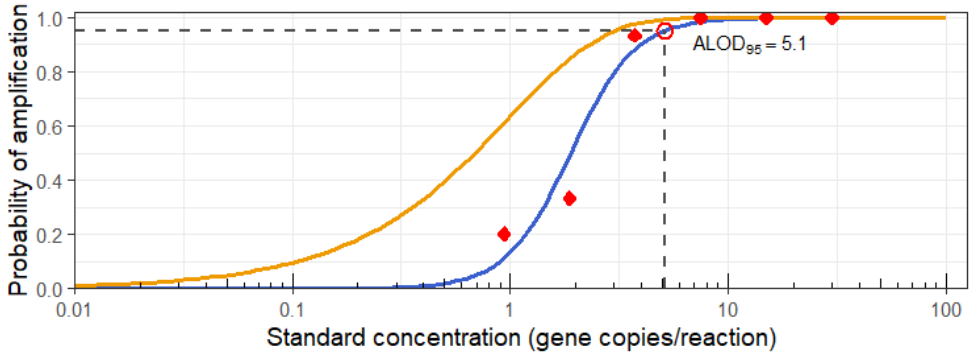


Figure 4.1. Empirical determination of the $ALOD_{95\%}$ by logistic regression. The observed proportion of samples at each dilution is denoted by the red diamonds; the blue curve is the fitted logistic model. The orange curve represents the theoretical maximum amplification possible which assumes consistently successful amplification of the standard material in which the number of targets is Poisson-distributed. The red circle denotes the estimated $ALOD_{95\%}$. Note that the orange curve at probability of amplification of 95% coincides with the theoretical 3 gene copies/reaction.

Based on the (binomial) logistic regression performed (Table 4.3, Figure 4.1), a standard concentration of 5.1 gene copies/reaction is expected to yield 95% probability of amplification. Assuming a concentration factor of 80× and an effective sample volume of 0.4 mL (e.g., Box 2.1), the $ALOD_{95\%}$ readjusted to the original processed sample volume would be reported as the $SLOD_{95\%}$ of 12.75 gc/mL.

The fitted logistic regression curve is shifted to the right compared to the theoretical maximum. The lower proportion of successful reactions (higher proportion of failed reactions) observed than that delineated by the theoretical maximum curve suggests that either not all targets present successfully amplified, or the number of targets present in each reaction were not Poisson distributed relative to the targeted mean concentrations and/or the standard material might have been of lower concentration than expected. This example reemphasizes that random sampling error of discrete targets in qPCR precludes the use of the chemical definition (and interpretation) of the ALOD as a direct indicator of assay sensitivity.

Method sensitivity

- "Method limit of detection" (MLOD) is defined as the "minimum concentration of a target necessary for its consistent detection after incorporating loss through the entire process from sample concentration to extraction" (Ahmed et al., 2021). The threshold for consistency is usually set to at least 95% detection.
- The MLOD requires rigorous and iterative testing to determine it (e.g., Derx et al., 2021). An approach to determine this value is described by Stokdyk et al. (2016). It involves spiking a range of concentrations of the target into each procedural step working backwards (i.e., extraction, secondary concentration, primary concentration) to determine the lowest concentration of the target in the sample matrix that would consistently yield a detection at the qPCR quantification step. This approach was applied using a recovery surrogate, which necessarily presumes that the surrogate spike seeded is representative of the recovery of the target.
- While potentially useful to compare different analytical methods and protocols, the MLOD is likely also sensitive to the recovery surrogate used and the sample matrix encountered, and would not be practical to evaluate for every method and sample matrix (Chik et al., 2018; Derx et al., 2021). Some aspects of method sensitivity might be captured by use of split samples spanning a gradient of low, moderate, and high anticipated target concentrations (regardless of whether targets were spiked or *in-situ*); however, caution should be exercised in broader extrapolation of the results beyond the range of conditions captured by this type of an approach.
- TAKE HOME MESSAGE FOR LABS: Given that the ability to *avoid* non-detect observations from a non-zero source (and therefore yield a detection) is a function of the original analytical process volume, the effective sample volume (and the associated concentration factor), and the surrogate analytical recovery profile (which provides an estimate of method losses throughout processing), this information should be reported in absence of a rigorously and iteratively determined MLOD.

5. Assay specificity

False-positives of targets can arise if primers and probes are not adequately specific to the targets of interest. BLAST searches are often performed *in silico* during the initial phases of primer/probe set development to identify target sequences that are adequately unique (i.e., specific) to the target for quantification. Cross-reactivity tests should be conducted during the initial phases of method development or when new primer/probes are being introduced to the assay (as may be the case when variants or new targets are introduced). The following examples of cross-reactivity related to wastewater surveillance assays for SARS-CoV-2 (or its surrogates/fecal indicators) have been previously noted; other cross-reactions might occur if primers and probes are not properly designed or selected:

- Certain SARS-CoV-2 RT-qPCR assays targeting the RdRP or M-genes with other endemic human coronaviruses (Westhaus et al., 2021)
- E_Sarbeco assay and seeded bovine coronavirus (Gerrity et al., 2020)
- Human coronavirus OC43 and seeded bovine coronavirus (Pecson et al., 2021)

The respective references for specific sequences that have been documented to result in cross-reactions should be consulted. To facilitate comparisons of assay specificity to the intended targets, the specific sequences used for primer/probes should be reported along with information pertaining to the RT-qPCR reaction (e.g., reaction volumes and conditions, master mix used).

6. Precision

Precision is the closeness of measurements to each other. This property can be evaluated based on the RT-qPCR assay using standard materials ("assay precision"), **OR** based on targets in real wastewater matrices considering method losses ("method precision"). Precision can also be determined within an assay (i.e., short-term precision during a single qPCR run, "intra-assay precision"; "repeatability"), or across multiple assays (i.e., long-term precision over multiple qPCR runs, "inter-assay precision"; "reproducibility"). Precision can only be determined for quantitative applications of RT-qPCR. Options for establishing limits for acceptable repeatability and reproducibility are summarized in Table 6.1.

Table 6.1. Options for establishing acceptable RT-qPCR repeatability and reproducibility control limits

Sample type analyzed	***Repeatability***	***Reproducibility***
Standard quantitation material	If technical replicates (n=2): • The absolute difference between duplicates (|ΔCq|) should be less than 0.8 Cq If technical replicates (n≥3): • Standard deviation of <0.5 Cq across technical triplicates Other equally or more stringent control* limits as established by control charting may be implemented	Mean Cq of technical replicates for a specific run observed within margin of error* established by control charting of long-term mean Cq's
Positive control (e.g., reference wastewater sample with SARS-CoV-2)		
Sample of interest	For routine monitoring using technical triplicates: • Standard deviation of <0.5 Cq across technical triplicates If target is expected to be present at low "trace" concentrations, greater variability of Cq values can be expected amongst replicates at lower target concentrations. Analysis of the sample by increasing effective sample size (e.g., more template volume or more technical replicates) is recommended.	N/A**

**If confidence intervals/margin of error on the mean are established to determine control limits, 95% confidence should be applied*

***Not typically determined*

Repeatability (Intra-assay precision)

- The precision of the assay with the same samples (regardless of whether standard quantitation materials, positive controls, or actual wastewater samples) repeatedly analyzed in the same assay is referred to as repeatability (short term precision; intra-assay variability).
- A minimum of three technical replicates are necessary for the characterization of repeatability through estimating standard deviations. Karlen et al. (2007) performed 144 PCR reactions at a range of concentrations based on the use of four or five replicate PCR reactions, and deemed replicates with standard deviations of less than 0.4 quantification cycles (Cq) across replicates of the same sample to be regarded to have acceptable repeatability; however, it was noted that the standard deviation of replicate Cq's increases with higher Cq values, with SD values less than 0.2 Cq for Cq's less than 30 cycles that can increase to over 0.8 Cq for Cq's over 30 cycles. Greater variability of Cq values can be expected amongst replicates at lower target concentrations than at higher target

concentrations because of the pronounced impact of random sampling error associated with small numbers of discrete targets (refer to Section on *Assay Sensitivity*).

- Various other works and protocols (e.g., *Institute of research in immunology and cancer, University of Montreal*, n.d.) suggest that a standard deviation of less than 0.5 Cq across technical triplicates for the same sample is generally regarded as acceptable repeatability (typically at concentrations anticipated to yield ≳10 gene copies/reaction).
    - If only technical duplicates (n=2) are performed, the absolute difference |ΔCq| between the duplicates (for standard quantitation materials or positive controls) can be monitored and should be less than 0.8 Cq. This is equally as stringent as a repeatability threshold of SD<0.5 Cq when n=3.
- Increasing the number of technical replicates can improve the concentration estimation when low levels of target are anticipated in samples (Taylor et al., 2019); however, in some circumstances this may not be feasible or possible (e.g., insufficient sample for analysis). In the absence of additional technical replicates, a review of historical standard curve data from the same method/instrument—especially those that correspond to the lowest standard dilution(s) prepared—might be useful to inform the degree of variability expected (and therefore permissible) at low target concentrations.
- The evaluation of repeatability of *standard quantitation materials and/or positive controls* can be more formally incorporated as part of a routine control charting process, which can then be used to establish lab- and instrument- specific warning and control limits (refer to section on *Control Charting*). Excessive variability amongst technical replicates of multiple *samples of interest* within a single run may also suggest the need for troubleshooting and/or a re-run.

Reproducibility (Inter-assay precision)

- Reproducibility (i.e., the long-term inter-assay precision) is usually evaluated based on the quantified standard material used to construct the standard curve or a positive control (e.g., wastewater known to yield detections for the target of interest). The average Cq (i.e., arithmetic mean) calculated for technical replicates of a quantitative positive control or specific dilution of the quantified standard material is expected to be relatively consistent between experiments. If the run-specific mean Cq lies beyond control limits (i.e., an outlier) as established through control charting OR exhibits systematic bias, the results of the experiment are highly suspect (refer to following section on *Control Charting*).

Control charting

Control charting (statistical landmarks overlaid on a plot) is an effective means to monitor both intra- and inter-assay variability and can be used to complement other QA/QC measures taken. The variability exhibited at any level of a standard curve's dynamic range can be tracked. This can help prevent deterioration of a measurement process, provide diagnostic information, demonstrate the degree of repeatability and reproducibility attained, and to flag run-specific issues (e.g., instrument/reagent signal drift; Laboratory Services Branch, n.d.; J. K. Taylor, 1987). Control charting can also be used to establish key warning and control thresholds (Box 6.1).

- <u>TAKE HOME MESSAGE FOR LABS:</u> Tracking of method performance and establishing warning and control limits is an essential part of QA/QC of results emanating from RT-qPCR. Control charting can be adopted and formalized to track and improve method development and optimization. Laboratory-specific rules can be established with the help of a statistician to provide confidence that the assay is performing within its expected operational conditions (example in Box 6.1).

Box 6.1 Use of control charts to monitor intra-assay variability (repeatability) and inter-assay variability (reproducibility)

Control charts can be developed for each dilution of the standard curve. For a laboratory that runs technical duplicates, the absolute difference in Cq between the duplicates was tracked for repeatability (Figure 6.1); the ongoing mean difference is tracked, along with thresholds with warning and control limits. This facilitates real-time review of control charts. Rules can be established (Figure 6.2, see y-axis) to trigger further method troubleshooting/action. Similarly, the mean Cq from each qPCR run can also be tracked over time to track reproducibility (Figure 6.3), with analogous thresholds and rules determined to trigger method troubleshooting (MECP Procedures Manual LSBSOP041).

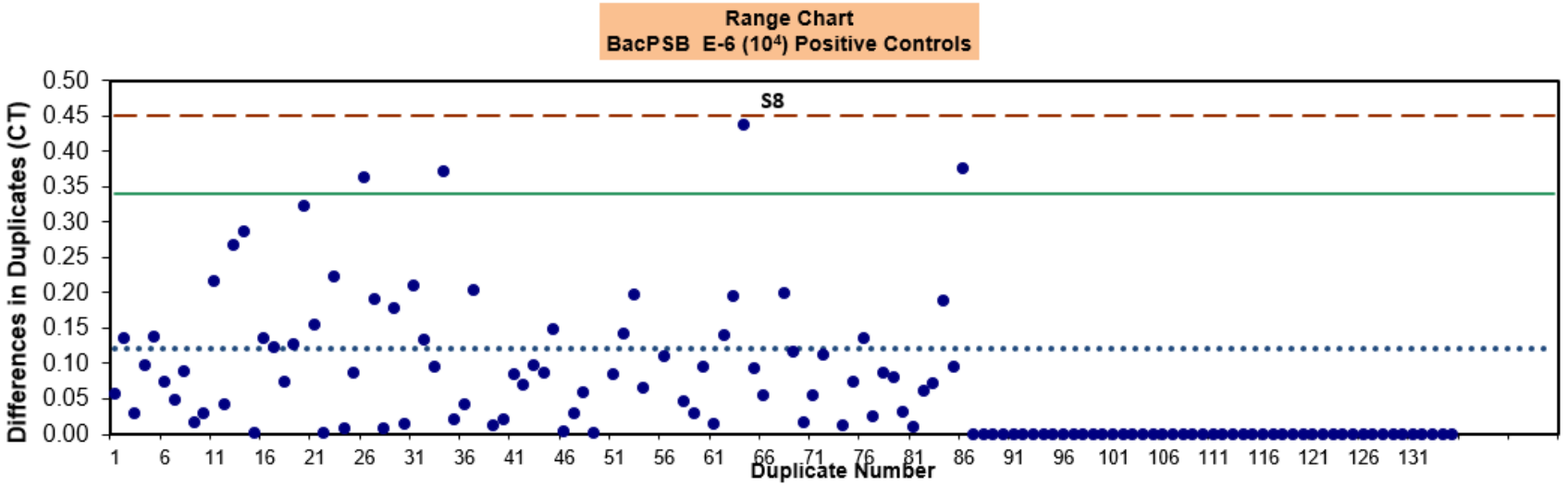


Figure 6.1. Tracking of absolute difference in Cq between the technical duplicates for repeatability

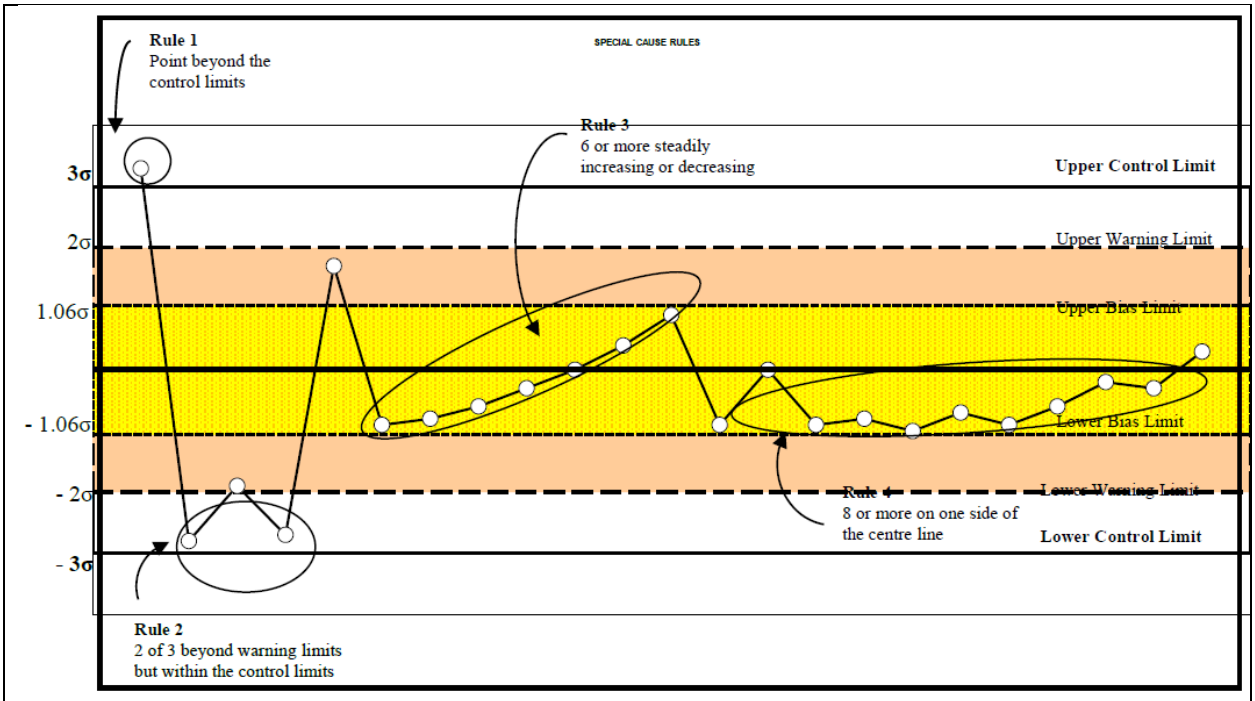


Figure 6.2. Rules established to trigger further method troubleshooting/action by control charting

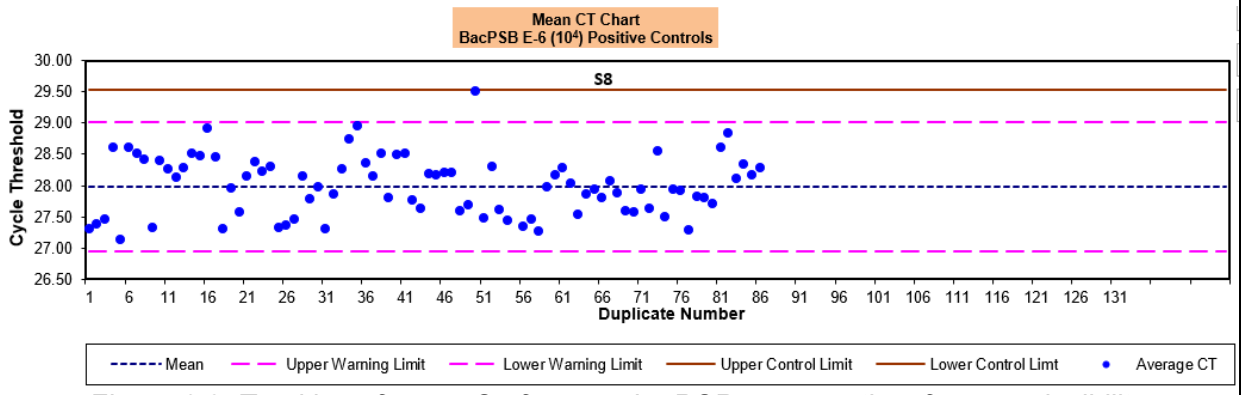


Figure 6.3. Tracking of mean Cq from each qPCR run over time for reproducibility

Statistical measures of assay precision

- The coefficient of variation (CV) is a statistical measure of the standard deviation relative to the mean of replicate observations at each dilution. This measure should not be calculated based on the Cq values directly; instead it must be related to the scale used for concentration estimation by accounting for the log-linear relationship between the Cq values and concentration estimates. Therefore, this important distinction has been denoted as CV_{ln} (Equation 3, Kralik & Ricchi, 2017):

$$CV_{ln} = \sqrt{(1+E)^{(SD(C_q))^2 \times ln(1+E)} - 1} \qquad \textit{(Equation 3)}$$

  Where CV_{ln} is the coefficient of variation

  E is the PCR efficiency consistent with the slope of the overall standard curve (E=$10^{-1/slope}$-1)

  $SD(C_q)$ is the standard deviation of repeated measurements of the cycle threshold corresponding to a specific concentration of a standard

- CV_{ln} values from 10% to 35% are commonly used as acceptance criteria (Forootan et al., 2017; Haugland et al., 2016; Klymus et al., 2020); some groups have even proposed fixing the CV_{ln} at 25% (Kralik & Ricchi, 2017). However, it has been unclear whether these documented criteria have been consistently applied respective to repeatability or reproducibility.

- Usually, a more stringent threshold of ≤25% would be applied for repeatability than for reproducibility (≤35%) for the dynamic range of the standard curve. These thresholds are typically determined using a statistical sample size of at least 15 repeats per standard dilution (European Commission. Joint Research Centre. Institute for Health and Consumer Protection, 2008)

- The “assay limit of quantification” (ALOQ) has been defined as the lowest concentration of a standard material below which a specified level of accuracy/precision, based on the CV_{ln}, would not be satisfied. The quantitative value of this threshold will necessarily be greater than sensitivity-related thresholds (e.g., ALOD) for a given assay because precision can only be determined once a substance has been ascertained to be present.

- There is broad alignment amongst practitioners that the ALOQ qualitatively corresponds to the point of inflection on a standard curve capturing a wide dynamic range that divides the linear and non-linear regions (and quantification in the non-linear region has been deemed less reliable). The majority of documented approaches used to determine the ALOQ are based on the coefficient of variation calculated for technical replicates of each standard material dilution. When this is evaluated in a single experiment, the ALOQ is a measure of precision related to

repeatability. When this is evaluated over multiple experiments, the ALOQ is reflective of reproducibility.

- The concentrations used in the standard curve which exceed the specified levels of precision have been directly regarded as the ALOQ (for example, an assay in which the 100 gene copies standard has a CV_{ln} of 15%, and the 10 gene copies standard has a CV_{ln} of 75% would have an $ALOQ_{15\%}$ of 100 gene copies); however, it is a more common practice that $\%CV_{ln}$ is graphically plotted against log(standard concentration), and the point at which the CV_{ln} threshold is exceeded would be regarded the ALOQ (Box 6.2).

- TAKE HOME MESSAGES FOR LABS: An interim target of CV_{ln}=35% will be used for the determination of the ALOQ (regardless of repeatability or reproducibility). It is recommended that the practice of graphically plotting $\%CV_{ln}$ against log (standard concentration) be adopted to determine the $ALOQ_{35\%}$.

- The evaluation of experiment-specific ALOQ for *every* qPCR run with 15 technical replicates at each dilution spanning the dynamic range is impractical. Accordingly, the ALOQ associated with *repeatability* ($ALOQ_r$) corresponding to a CV_{ln} of 35% shall be evaluated using data collected from the same experiment performed for the determination of $ALOD_{95\%}$ (refer to section on *Assay Sensitivity* and Box 4.1).

- An examination of datasets provided by some participants of Ontario's WSI revealed the need for stringent and consistent pre-handling of standard curves with established acceptance criteria prior to determination of an ALOQ related to *reproducibility* ($ALOQ_R$). Labs participating within the WSI are asked to determine the $ALOQ_R$ using standard curves constructed over 30 experiments. All labs should set the highest dilution (i.e., lowest concentration) used for the standard curve at an intended mean target concentration at which no "drop outs" occur (refer to section on *Sensitivity*). For these 30 experiments, labs are asked to quantify technical triplicates of an additional dilution of standard material (at approximately half of the lowest concentration used in the standard dilution series), *but not used for the quantification based on the standard curve*. This point will be monitored for the proportion of "drop-outs" (see Box 4.1).

- A companion spreadsheet to this *Protocol* has been developed to support the determination of the $ALOQ_r$ (and $ALOD_{95\%}$) as described above. The same spreadsheet can be used in the calculation of the $ALOQ_R$ with data compiled over multiple qPCR runs. Both $ALOQ_r$ and $ALOQ_R$ will need to be readjusted using the CF or ESV and reported as the $SLOQ_r$ and $SLOQ_R$, respectively.

- **As an *interim* measure for consistent reporting and interpretation, an observed sample Cq falling between the Cq associated with the lowest point**

of the standard curve (at approximately 5 gene copies/reaction) and the y-intercept of the experiment-specific standard curve will be reported as an extrapolated concentration estimate amended with a "J" qualifier. If a Cq beyond the y-intercept is observed but clear amplification was observed, the qualifier "UJ" should be used instead to denote a "trace" of signal.

- **The above *interim* measure will be used for reporting in the WSI until sufficient data has been generated and the utility (practicality and statistical implications) of these statistical measures ($SLOD_{95\%}$, $SLOQ_r$, and $SLOQ_R$) can be reviewed and confirmed.**

Method precision

- "Method limit of quantification" (MLOQ) is defined as "the lowest concentration of a target that can be quantified with an acceptable level of precision when present in a sample" (Ahmed et al., 2021).
- The determination of this value would require an approach analogous to the determination of the MLOD, but would suffer from similar limitations to its determination and ultimately, its usefulness (i.e., matrix-, surrogate-, and method-specific).
- Some aspects of method precision might be captured by use of split samples spanning a gradient of high to low anticipated target concentrations (regardless of whether targets were spiked or *in-situ*); however, caution should be exercised in broader extrapolation of the results beyond the range of conditions captured by this type of an approach.

Box 6.2 Evaluation of $SLOQ_r$ (SLOQ-repeatability) and $SLOQ_R$ (SLOQ-reproducibility)
The repeatability-based $ALOQ_r$ (and its associated $SLOQ_r$) is determined using the observed Cq's corresponding to the dataset acquired to determine the $ALOD_{95\%}$ (Box 4.1). For the 7.5 gene copies/reaction dilution in Box 4.1, the following Cq's were observed:

35.47	35.93	34.54	35.26	36.01
35.30	34.56	36.42	35.36	36.02
35.87	35.52	36.80	35.60	34.93

Given that these Cq's were subject to control limits established for the method, the PCR efficiency can be assumed to be consistent with that corresponding to the long-term standard curve.
Slope: -3.393, Intercept: 38.348, RSQ=0.9985
PCR efficiency: 0.97

Sample calculation:
Mean Cq 35.448
SD Cq 0.630

$$CV_{ln} = \sqrt{(1+E)^{(SD(C_q))^2 \times ln(1+E)} - 1} \qquad \textit{(Equation 3)}$$

$$CV_{ln} = \sqrt{(1+0.97)^{(0.63)^2 \times ln(1+0.97)} - 1}$$

$$CV_{ln} = 44.8\%$$

Suppose the CV_{ln}(%) corresponding to 15 gene copies/reaction and 30 gene copies/reaction were calculated to be 32.7% and 22.9%, respectively. Simple linear interpolation of CV_{ln} of 35% yields an $ALOQ_r$ of ~13.6 gene copies/reaction. Assuming a concentration factor of 80× and effective sample volume of 0.4 mL (e.g., Box 2.1), the $ALOQ_r$ readjusted to the original processed sample volume would be reported as the $SLOQ_r$ of 34 gene copies/mL.

Similarly, the $SLOQ_R$ can be evaluated using the mean Cq values and the standard deviation *of the mean Cq values* for each dilution used in standard curves across all experiments.

N.B. (1) The estimation of CV_{ln} will be biased when a dilution that does not always yield amplification of signal is used. Therefore, calculation of CV_{ln} should only be performed on standard dilutions for which drop-outs of signal do not occur.

N.B. (2) At present, there is insufficient empirical evidence gathered to suggest the broad applicability and utility of the $SLOQ_r$ and $SLOQ_R$. Circumstances may arise when the CV_{ln} does not decrease monotonically with increasing standard concentration, or if the CV_{ln} is consistently low (e.g., <20%). Accordingly, sufficient data is needed to verify the full utility of the proposed approach.

7. Analytical recovery

Analytical recovery of the method is intricately linked to its sensitivity. A known amount of a recovery surrogate (that is deemed to represent the behaviour of the target of interest) is typically seeded into the analytical sample portion as a positive control and subjected to the same experimental procedures as the target of interest to evaluate how much of the original surrogate can be successfully observed and captured by the final measurement. Despite the absence of a single, universal recovery surrogate of SARS-CoV-2, when a specific recovery surrogate is administered using a consistent protocol

into a common wastewater matrix, comparisons of analytical recovery amongst methods can be facilitated (Kantor et al., 2021).

Many model viruses and potential surrogates with similar structural and morphological characteristics to SARS-CoV-2 have been investigated for their ability to reflect SARS-CoV-2 behaviour in wastewater. These include:

- Human coronavirus OC43 (Pecson et al., 2021)
- HCoV-229E (Chik et al., 2021)
- Mouse Hepatitis Virus (MHV; Ahmed, Bertsch, Bivins, et al., 2020)
- Bovine coronavirus (LaTurner et al., 2021)
- Bovine respiratory syncytial virus (Gonzalez et al., 2020)
- Armored RNA (Gonzalez et al., 2020)
- Bacteriophage Phi6 (Ye et al., 2016)
- F-specific RNA phages (Medema et al., 2020)
- MS2 bacteriophage (Forés et al., 2021)
- Vesicular stomatitis virus (VSV; D'Aoust et al., 2020)
- Inactivated SARS-CoV-2 (Ahmed et al., 2020)

- It has been recognized that some recovery surrogates (e.g., inactivated SARS-CoV-2, human coronavirus OC43) are likely more appropriate to reflect the behaviour of SARS-CoV-2 in wastewater than others (e.g., F-specific RNA phages, MS2 bacteriophage).
- Caution should be exercised in broader extrapolation of the surrogate recovery results to that of *in-situ* SARS-CoV-2 present in various wastewater matrices. Of note, the extraction of the spiked surrogate is not likely to be 100% efficient and consistent with the extraction efficiency of the target of interest, therefore calculated recovery efficiencies are subject to bias (Kantor et al., 2021).
- Generally, recovery surrogates can be deployed at any step throughout the analytical procedure (Box 7.1). Whole process controls can be seeded into the samples before processing to account for losses throughout the entire procedure; extraction efficiency controls are administered prior to the RNA extraction step. In addition, these same process controls can be repeated in another blank matrix (e.g., water). No universal blank matrix exists; some methods require a blank matrix other than water.
- The administration of the recovery surrogate in various ways have been used to quantify the following recovery ratios ("RR"):
    - Whole process recovery ("analytical recovery"): The ratio of surrogate recovered from the wastewater sample as the whole process control

(numerator) *relative to* the surrogate recovered from the extraction efficiency control without the wastewater sample matrix (denominator).

- Extraction efficiency (also referred to as "RNA/DNA isolation efficiency"): The ratio of extraction efficiency control recovered from the wastewater sample (numerator) *relative to* that same control subjected to the same extraction process without the wastewater sample matrix (denominator).
- Pre-treatment step recovery: The ratio of whole process control recovered (numerator) *relative to* the extraction efficiency control recovered in the wastewater sample (denominator).

- Caution should be exercised to determine whether the recovery surrogates are deployed in such a fashion that necessitates handling concentration estimates in the calculation of the recovery ratio as paired data. For example, when six aliquots (i.e., biological replicates labelled A through F) of the same sample are prepared, with three of the aliquots (e.g., aliquot A, B, C) spiked with a recovery surrogate as whole process controls and the others spiked at the extraction efficiency step (e.g., aliquot D, E, F), the aliquots are <u>not</u> paired (aliquot A is <u>not</u> paired with aliquot D). Rather, the numerator and denominator of the various recovery estimates are evaluated as separate, independent replicates (i.e., the same surrogate will not be spiked at both Spike Point A and B within the same aliquot; Figure 7.1). Accordingly, the estimation of mean and variance of the recovery ratios ("RR") may be estimated by the approximations:

$$mean(RR) \approx \frac{\mu_x}{\mu_z}\left(1 + \frac{\sigma_z^2}{\mu_z^2}\right) \qquad \textit{(Equation 4)}$$

$$variance(RR) \approx \frac{1}{\mu_z^2}\left(\sigma_x^2 + \frac{\mu_x^2}{\mu_z^2}\sigma_z^2\right) \qquad \textit{(Equation 5)}$$

where the subscripts *x* and *z* denote the numerator and denominator of the recovery ratio of interest, and μ and σ represent the mean and standard deviation, respectively

- For the purpose of methods development or methods comparison (such as an inter-laboratory study designed to compare methods), the determination of all three aspects of recovery should be quantified to track losses.
- For the purpose of ongoing monitoring and data reporting, the "overall" surrogate whole process recovery should be reported with analytical results of the target of interest to contextualize the data.

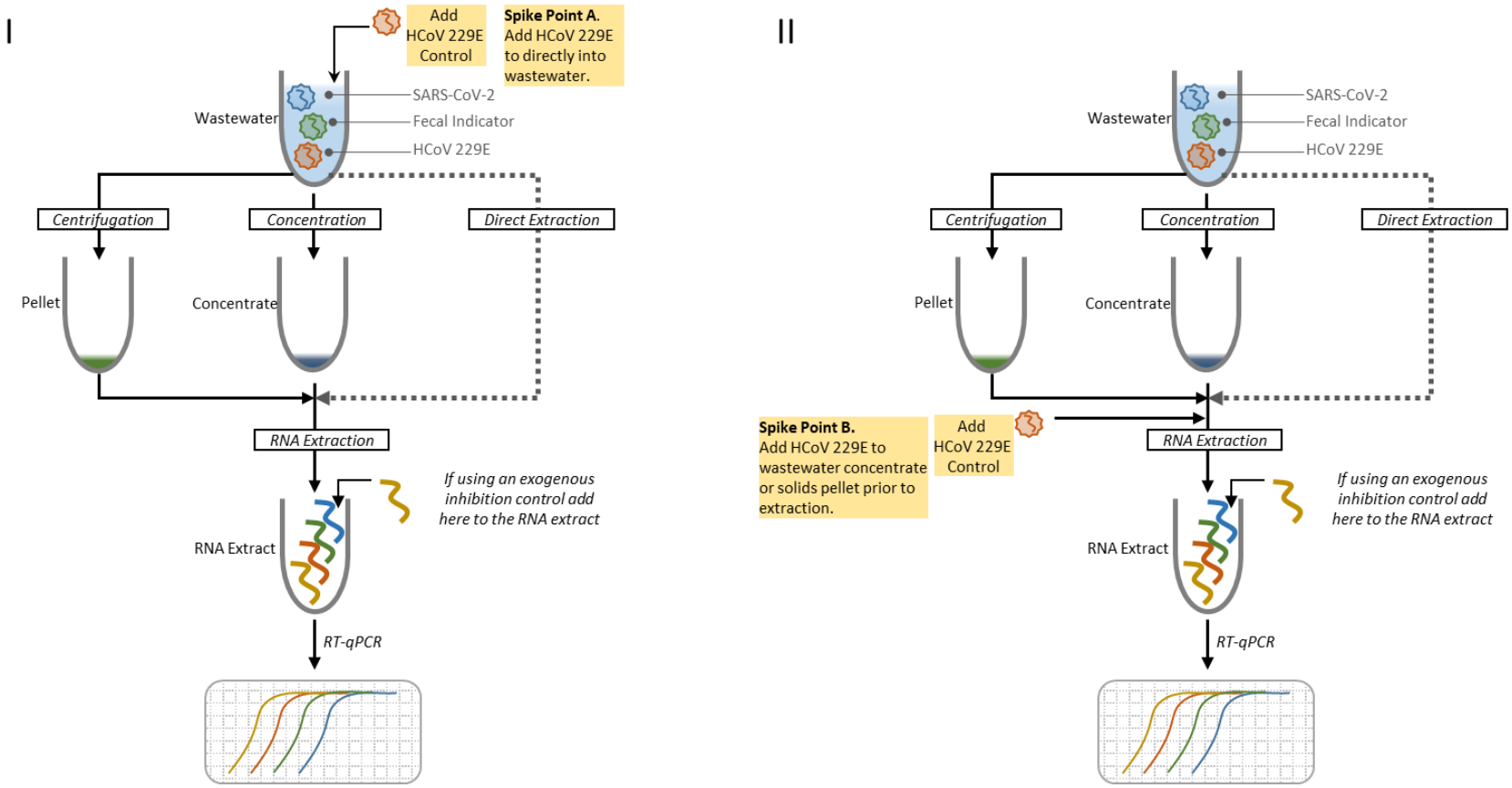


Figure 7.1. The administration of recovery surrogates to a wastewater sample. In this example, human coronavirus strain 229E (HCoV-229E) has been propagated and introduced as the recovery surrogate. After aliquoting the sample to the desired sample process volume, the recovery surrogate can be administered at spike point “A” (I) or spike point “B” (II) in separate aliquots of the wastewater sample. This can also be repeated in a blank matrix, although little value has been shown in administering a recovery surrogate to spike point “A” in clean water and processed using some methods. For example, using some centrifugation-based methods, the lack of a solids matrix available for the preferential partitioning of the recovery surrogate might preclude its recovery.

8. Other QA/QC controls

The *Environmental Microbiology Minimum Information (EMMI) Guidelines* (Borchardt et al., 2021) provides an extensive list of positive and negative controls that can be deployed at various steps of sample handling and processing, along with guidance for reporting and interpretations. Negative controls are used to assess and troubleshoot contamination; positive controls are used to assess/troubleshoot recovery, efficiency and inhibition. The following sections provide minimum requirements and associated interpretations of specific controls consistent with the framework set out by this *Protocol.*

Failed negative controls

- Negative controls, such as whole processing blanks and no template controls (NTCs), can be used to provide assurance that no contamination has occurred. Ideally, every PCR run should not yield an amplification of signal in the negative controls; however, there are cases where a low level of background contamination does not substantially impact the interpretation of the results (e.g., when sample signal greatly exceeds that observed in the negative controls) and it may be impractical to perform a re-run. In cases where failed negative controls occurred but are not expected to impact interpretation of sample signals, a “B” qualifier in reporting should be used to document possible background contamination.
- The number of negative controls required to detect contamination can be theoretically determined for a specified probability of detection by assuming the contaminating gene target is Poisson distributed. For example, performing three NTCs will provide 95% probability of seeing contamination at a level of 1 copy of contaminant per reaction. In practice, negative controls replicated at multiple steps of sample handling and processing will provide confidence that contamination issues can be detected and addressed (Borchardt et al., 2021).
- Within the Ontario WSI, a minimum threshold for assigning a “B” qualifier has been proposed based on NTCs (Figure 1). For failed NTCs of any assay, the level of contamination observed should be less than the experiment-specific standard curve y-intercept; moreover, the level of amplification exhibited in any of the samples of interest must be $\gtrsim$5 Cq than observed in the failed NTC with the greatest degree of contamination. A re-run of the sample with appropriate controls is warranted for the target(s) for which the NTCs do not meet such criteria.
- Alternate forms of corrective action and troubleshooting may be needed for failure of other negative controls (e.g., failed whole process blanks without failed NTCs). The laboratory analyst will need to evaluate the degree to which samples on that experimental run may/may not have been compromised, and provide adequate documentation to justify the corrective actions taken.

Failed inhibition controls

Inhibition controls are an example of a type of positive control performed to ensure that sample concentration estimates have not been negatively impacted due to inhibitory substances present within the sample matrix. A range of inhibition controls can be performed, including the use of endogenous viruses within the sample, or controls seeded into the sample matrix. Strategies for troubleshooting and addressing inhibition have been extensively documented in other published works (e.g., Gibson et al., 2012) and therefore, will not be discussed in detail here.

- Samples deemed to be inhibited should be re-diluted and re-run. If inhibition has been successfully addressed through the re-run, an “AI” qualifier can be reported with the updated analytical result.
- While the dilution of endogenous targets might be useful to assess inhibition, the target of interest may already be present at low levels, or there may be insufficient sample available for additional testing. In these cases, the “FI” qualifier should be reported with these inhibited results.

9. Implications for reporting, analysis and interpretation

Accurate representation of the analytical results and documentation of QA/QC measures taken are necessary to facilitate the adoption of appropriate statistical approaches for data analysis. For RT-qPCR results estimated within the operational dynamic range of the assay, the following information shall be reported:

- the final concentration estimate of the target in the original sample;
- the original sample volume processed;
- the effective volume assayed; and
- an estimate of whole process analytical recovery based on a surrogate (if performed).

It is acknowledged that in the absence of a recovery surrogate that has been ascertained to behave like *in-situ* SARS-CoV-2 recovered from wastewater, the use of a single “best-available” surrogate is recommended. The choice of surrogate (e.g., HCoV-229E, MHV, armored RNA, etc.) should be reported along with the results of the recovery assay.

- For the purpose of unbiased concentration estimation, raw data of technical replicates are required. An appropriate statistical model that represents and accounts for the stochastic nature of sampling/subsampling/amplification/ quantitation should be applied.
- A non-detect shall be reported when no amplification of the target occurs. As the non-detect may reflect a failure to capture, concentrate, and/or amplify any targets within the reaction volume, it must not be interpreted or construed as though the

analyte is absent. It must also not be reported (or interpreted) as a concentration estimate of <1 gene copy per reaction.

- As an interim measure, observed sample Cq values falling between the experiment-specific standard curve's y-intercept and the lowest concentration used in the standard curve shall be reported as extrapolated concentration estimates denoted with a "J" qualifier.
- "Trace" detections (amplification of signal beyond the y-intercept of the experiment-specific standard curve) shall be denoted with a "UJ" qualifier. Such a positive test result is indicative of the presence of at least one target (≥1) in the reaction volume. When trace levels of the target are expected, it is prudent to analyze additional technical replicates and/or increase the effective sample size captured by the assay to allow for the estimation of concentration at these low levels.
- For the purpose of data visualization of trends—the substitution of placeholder values for non-detects and "trace" detects is commonly applied (e.g., ½ SLOD). However, the approach applied must be explicitly stated to allow for a transparent interpretation of the data. Similarly, slopes calculated for trend analysis using substituted values should also clearly indicate how non-detect and "trace" detects were manipulated, as doing so will likely incur bias.

Although all measurements are imperfect, there are circumstances where results emanating from a qPCR analysis are subject to greater bias (e.g., background contamination), variability and/or uncertainty (e.g., estimation of concentrations when low number of targets are present) than normally expected of the methodology deployed. In such cases, the use of qualifiers is warranted to facilitate the transparent articulation of these results, such that data analysts and those responsible for data interpretation handle the data with additional caution. The following qualifiers have been suggested:

- **B**: Analytical result may be subject to "trace" levels of contamination; the target analyte was also detected in negative controls (below the y-intercept of the experiment-specific standard curve, but at least 5 Cq lower than the Cq associated with the contaminated negative control, Figure 1) on the same run as the sample
- **AI:** The original sample was inhibited; however, the inhibition has been addressed through dilution. The reported concentration estimate is the updated result after addressing inhibition.
- **FI:** The original sample was inhibited; however, the inhibition has not been successfully addressed. Therefore, no concentration estimate has been reported.
- **ND:** no amplification occurred in the reaction; non-detect

- **J**: *[Interim usage]* Analytical result falls below the lowest concentration of the experiment-specific standard curve but above the y-intercept value; the reported value is based on the extrapolation of the standard curve in the non-linear region
- **UJ**: *[Interim usage]* Observed Cq is greater than the experiment-specific standard curve intercept value but evidence of clear amplification was present (i.e., "trace" signal observed); the reported value is based on the extrapolation of the standard curve in the non-linear region

Statistical measures of sensitivity and precision (i.e., $SLOD_{95\%}$, $SLOQ_r$, $SLOQ_R$) evaluated consistently with the terms of this *Protocol* should also be reported with analytical results. However, until the utility and statistical implications of these measures can be reviewed and confirmed, the reported values will not be used in the interpretation of the reported data.

Assuming that the lowest concentration of the standard curve corresponded to 5 gc/reaction, examples of these qualifiers are presented in Table 9.1:

Table 9.1 Example for the use and interpretation of interim qualifiers in the reporting of RT-qPCR results

Example	…	Result	Units	Sample volume (mL)	CF	ESV (mL)	Surrogate recovery estimate (%)	Recovery Surrogate	Qualifier
Cq within dynamic range		24.5	gc/mL	40	100	2	15.8	HCoV-229E	
Cq between lowest standard curve concentration and y-intercept		3.2	gc/mL	40	100	2	15.8	HCoV-229E	J
Cq>y-intercept (“trace”)		0.5	gc/mL	40	100	2	15.8	HCoV-229E	UJ
Non-detect		ND	gc/mL	40	100	2	15.8	HCoV-229E	ND
Separate run									
Failed NTC; Cq_{sample}-$Cq_{NTC} \gtrsim 5$		10.5	gc/mL	40	100	2	18.3	HCoV-229E	B
Inhibited sample		-	gc/mL	40	100	2	2.2	HCoV-229E	FI
Sample after inhibition addressed		75.6	gc/mL	40	20	0.4	13.4	HCoV-229E	AI

10. References and additional resources

Ahmed, W., Bertsch, P. M., Bibby, K., Haramoto, E., Hewitt, J., Huygens, F., Gyawali, P., Korajkic, A., Riddell, S., Sherchan, S. P., Simpson, S. L., Sirikanchana, K., Symonds, E. M., Verhagen, R., Vasan, S. S., Kitajima, M., & Bivins, A. (2020). Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology. *Environmental Research*, 110092. https://doi.org/10.1016/j.envres.2020.110092

Ahmed, W., Bertsch, P. M., Bivins, A., Bibby, K., Farkas, K., Gathercole, A., Haramoto, E., Gyawali, P., Korajkic, A., McMinn, B. R., Mueller, J. F., Simpson, S. L., Smith, W. J. M., Symonds, E. M., Thomas, K. V., Verhagen, R., & Kitajima, M. (2020). Comparison of virus concentration methods for the RT-qPCR-based recovery of murine hepatitis virus, a surrogate for SARS-CoV-2 from untreated wastewater. *Science of the Total Environment*, *739*, 139960. https://doi.org/10.1016/j.scitotenv.2020.139960

Ahmed, W., Simpson, S., Bertsch, P., Bibby, K., Bivins, A., Blackall, L., Bofill-Mas, S., Bosch, A., Brandao, J., Choi, P., Ciesielski, M., Donner, E., D'Souza, N., Farnleitner, A., Gerrity, D., Gonzalez, R., Griffith, J., Gyawali, P., Haas, C., … Shanks, O. (2021). *Minimizing errors in RT-PCR detection and quantification of SARS-CoV-2 RNA for wastewater surveillance*. https://doi.org/10.20944/preprints202104.0481.v1

Bivins, A., Kaya, D., Bibby, K., Simpson, S., Bustin, S., Shanks, O., & Ahmed, W. (2021). *Inherent bias of SARS-CoV-2 RNA quantification for wastewater surveillance due to variable RT-qPCR assay parameters*. https://doi.org/10.20944/preprints202106.0320.v1

Borchardt, M. A., Boehm, A. B., Salit, M., Spencer, S. K., Wigginton, K. R., & Noble, R. T. (2021). The Environmental Microbiology Minimum Information (EMMI) guidelines: qPCR and dPCR quality and reporting for Environmental Microbiology. *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.1c01767

Broeders, S., Huber, I., Grohmann, L., Berben, G., Taverniers, I., Mazzara, M., Roosens, N., & Morisset, D. (2014). Guidelines for validation of qualitative real-time PCR methods. *Trends in Food Science & Technology*, *37*(2), 115–126. https://doi.org/10.1016/j.tifs.2014.03.008

Burns, M., & Valdivia, H. (2008). Modelling the limit of detection in real-time quantitative PCR. *European Food Research and Technology*, *226*(6), 1513–1524. https://doi.org/10.1007/s00217-007-0683-z

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., & Wittwer, C. T. (2009). The MIQE Guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, *55*(4), 611–622. https://doi.org/10.1373/clinchem.2008.112797

Chik, A. H.S., Glier, M. B., Servos, M., Mangat, C. S., Pang, X.-L., Qiu, Y., D'Aoust, P. M., Burnet, J.-B., Delatolla, R., & Dorner, S. (2021). Comparison of approaches to quantify SARS-CoV-2 in wastewater using RT-qPCR: Results and implications from a collaborative inter-laboratory study in Canada. *Journal of Environmental Sciences*, *107*, 218–229.

Chik, A. H. S., Schmidt, P. J., & Emelko, M. B. (2018). Learning something from nothing: The critical importance of rethinking microbial non-detects. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.02304

D'Aoust, P. M., Mercier, E., Montpetit, D., Jia, J.-J., Alexandrov, I., Neault, N., Baig, A. T., Mayne, J., Zhang, X., Alain, T., Servos, M. R., MacKenzie, M., Figeys, D., MacKenzie, A. E., Graber, T. E., & Delatolla, R. (2020). Quantitative analysis of SARS-CoV-2 RNA

from wastewater solids in communities with low COVID-19 incidence and prevalence. *MedRxiv*, 2020.08.11.20173062. https://doi.org/10.1101/2020.08.11.20173062

Derx, J., Demeter, K., Linke, R., Cervero-Aragó, S., Lindner, G., Stalder, G., Schijven, J., Sommer, R., Walochnik, J., Kirschner, A. K. T., Komma, J., Blaschke, A. P., & Farnleitner, A. H. (2021). Genetic Microbial Source Tracking support QMRA modeling for a riverine wetland drinking water resource. *Frontiers in Microbiology*, *0*. https://doi.org/10.3389/fmicb.2021.668778

European Commission. Joint Research Centre. Institute for Health and Consumer Protection. (2008). *Definition of minimum performance requirements for analytical methods of GMO testing: European Network of GMO Laboratories (ENGL)*. Publications Office. https://data.europa.eu/doi/10.2788/65827

Forés, E., Bofill-Mas, S., Itarte, M., Martínez-Puchol, S., Hundesa, A., Calvo, M., Borrego, C. M., Corominas, L. L., Girones, R., & Rusiñol, M. (2021). Evaluation of two rapid ultrafiltration-based methods for SARS-CoV-2 concentration from wastewater. *Science of the Total Environment*, *768*, 144786. https://doi.org/10.1016/j.scitotenv.2020.144786

Forootan, A., Sjöback, R., Björkman, J., Sjögreen, B., Linz, L., & Kubista, M. (2017). Methods to determine limit of detection and limit of quantification in quantitative real-time PCR (qPCR). *Biomolecular Detection and Quantification*, *12*, 1–6. https://doi.org/10.1016/j.bdq.2017.04.001

Gerrity, D., Papp, K., Stoker, M., Sims, A., & Frehner, W. (2020). Early-pandemic wastewater surveillance of SARS-CoV-2 in Southern Nevada: Methodology, occurrence, and incidence/prevalence considerations. *Water Research X*, *10*. https://doi.org/10.1016/j.wroa.2020.100086

Gibson, K. E., Schwab, K. J., Spencer, S. K., & Borchardt, M. A. (2012). Measuring and mitigating inhibition during quantitative real time PCR analysis of viral nucleic acid extracts from large-volume environmental water samples. *Water Research*, *46*(13), 4281–4291. https://doi.org/10.1016/j.watres.2012.04.030

Gonzalez, R., Curtis, K., Bivins, A., Bibby, K., Weir, M. H., Yetka, K., Thompson, H., Keeling, D., Mitchell, J., & Gonzalez, D. (2020). COVID-19 surveillance in Southeastern Virginia using wastewater-based epidemiology. *Water Research*, *186*, 116296. https://doi.org/10.1016/j.watres.2020.116296

Haugland, R. A., Siefring, S., Varma, M., Oshima, K. H., Sivaganesan, M., Cao, Y., Raith, M., Griffith, J., Weisberg, S. B., Noble, R. T., Blackwood, A. D., Kinzelman, J., Anan'eva, T., Bushon, R. N., Stelzer, E. A., Harwood, V. J., Gordon, K. V., & Sinigalliano, C. (2016). Multi-laboratory survey of qPCR enterococci analysis method performance in U.S. coastal and inland surface waters. *Journal of Microbiological Methods*, *123*, 114–125. https://doi.org/10.1016/j.mimet.2016.01.017

Institute of Research in Immunology and Cancer, University of Montreal. (n.d.). *Information on qPCR results* (Genomic Platform). Institute of Research in Immunology and Cancer, University of Montreal. https://genomique.iric.ca/resources/files/How_to_deal_with_qPCR_results.pdf

International Organization for Standardization. (2019). *ISO 20395:2019 Biotechnology—Requirements for evaluating the performance of quantification methods for nucleic acid target sequences—QPCR and dPCR*. https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/78/67893.html

Kantor, R. S., Nelson, K. L., Greenwald, H. D., & Kennedy, L. C. (2021). Challenges in measuring the recovery of SARS-CoV-2 from wastewater. *Environmental Science & Technology*, *55*(6), 3514–3519. https://doi.org/10.1021/acs.est.0c08210

Karlen, Y., McNair, A., Perseguers, S., Mazza, C., & Mermod, N. (2007). Statistical significance of quantitative PCR. *BMC Bioinformatics*, *8*, 131. https://doi.org/10.1186/1471-2105-8-131

Klymus, K. E., Merkes, C. M., Allison, M. J., Goldberg, C. S., Helbing, C. C., Hunter, M. E., Jackson, C. A., Lance, R. F., Mangan, A. M., Monroe, E. M., Piaggio, A. J., Stokdyk, J. P., Wilson, C. C., & Richter, C. A. (2020). Reporting the limits of detection and quantification for environmental DNA assays. *Environmental DNA*, *2*(3), 271–282. https://doi.org/10.1002/edn3.29

Kralik, P., & Ricchi, M. (2017). A basic guide to real time PCR in microbial diagnostics: Definitions, parameters, and everything. *Frontiers in Microbiology*, *8*. https://doi.org/10.3389/fmicb.2017.00108

Kralik, P., Slana, I., Kralova, A., Babak, V., Whitlock, R. H., & Pavlik, I. (2011). Development of a predictive model for detection of *Mycobacterium avium* subsp. Paratuberculosis in faeces by quantitative real time PCR. *Veterinary Microbiology*, *149*(1), 133–138. https://doi.org/10.1016/j.vetmic.2010.10.009

Laboratory Services Branch. (n.d.). *Procedure for Control Charting and Trend Analysis* (No. LSBSOP041, Revision 3). The Ministry of the Environment, Conservation and Parks.

LaTurner, Z. W., Zong, D. M., Kalvapalle, P., Gamas, K. R., Terwilliger, A., Crosby, T., Ali, P., Avadhanula, V., Santos, H. H., Weesner, K., Hopkins, L., Piedra, P. A., Maresso, A. W., & Stadler, L. B. (2021). Evaluating recovery, cost, and throughput of different concentration methods for SARS-CoV-2 wastewater-based epidemiology. *Water Research*, *197*, 117043. https://doi.org/10.1016/j.watres.2021.117043

Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., & Brouwer, A. (2020). Presence of SARS-coronavirus-2 RNA in sewage and correlation with reported COVID-19 prevalence in the early stage of the epidemic in The Netherlands. *Environmental Science & Technology Letters*, *7*(7), 511–516. https://doi.org/10.1021/acs.estlett.0c00357

Nolan, T., Huggett, J., Sanchez, E., Sanders, R., Redshaw, N., & Wilkes, T. (2013). *Good practice guide for the application of quantitative PCR (qPCR)*. https://doi.org/10.13140/RG.2.2.15943.96162

Nutz, S., Döll, K., & Karlovsky, P. (2011). Determination of the LOQ in real-time PCR by receiver operating characteristic curve analysis: Application to qPCR assays for *Fusarium verticillioides* and *F. proliferatum*. *Analytical and Bioanalytical Chemistry*, *401*(2), 717–726. https://doi.org/10.1007/s00216-011-5089-x

Pavšič, J., Devonshire, A. S., Parkes, H., Schimmel, H., Foy, C. A., Karczmarczyk, M., Gutiérrez-Aguirre, I., Honeyborne, I., Huggett, J. F., McHugh, T. D., Milavec, M., Zeichhardt, H., & Žel, J. (2015). Standardization of nucleic acid tests for clinical measurements of bacteria and viruses. *Journal of Clinical Microbiology*, *53*(7), 2008–2014. https://doi.org/10.1128/JCM.02136-14

Pecson, B. M., Darby, E., Haas, C. N., Amha, Y. M., Bartolo, M., Danielson, R., Dearborn, Y., Giovanni, G. D., Ferguson, C., Fevig, S., Gaddis, E., Gray, D., Lukasik, G., Mull, B., Olivas, L., Olivieri, A., Qu, Y., & Consortium, S.-C.-2 I. (2021). Reproducibility and sensitivity of 36 methods to quantify the SARS-CoV-2 genetic signal in raw wastewater: Findings from an interlaboratory methods evaluation in the U.S. *Environmental Science: Water Research & Technology*, *7*(3), 504–520. https://doi.org/10.1039/D0EW00946F

Ricchi, M., Savi, R., Bolzoni, L., Pongolini, S., Grant, I. R., Cicco, C. D., Cerutti, G., Cammi, G., Garbarino, C. A., & Arrigoni, N. (2016). Estimation of *Mycobacterium avium* subsp. Paratuberculosis load in raw bulk tank milk in Emilia-Romagna Region (Italy) by qPCR. *MicrobiologyOpen*, *5*(4), 551–559. https://doi.org/10.1002/mbo3.350

Slana, I., Kralik, P., Kralova, A., & Pavlik, I. (2008). On-farm spread of *Mycobacterium avium* subsp. Paratuberculosis in raw milk studied by IS900 and F57 competitive real time

quantitative PCR and culture examination. *International Journal of Food Microbiology*, *128*(2), 250–257. https://doi.org/10.1016/j.ijfoodmicro.2008.08.013

Stokdyk, J. P., Firnstahl, A. D., Spencer, S. K., Burch, T. R., & Borchardt, M. A. (2016). Determining the 95% limit of detection for waterborne pathogen analyses from primary concentration to qPCR. *Water Research*, *96*, 105–113. https://doi.org/10.1016/j.watres.2016.03.026

Svec, D., Tichopad, A., Novosadova, V., Pfaffl, M. W., & Kubista, M. (2015). How good is a PCR efficiency estimate: Recommendations for precise and robust qPCR efficiency assessments. *Biomolecular Detection and Quantification*, *3*, 9–16. https://doi.org/10.1016/j.bdq.2015.01.005

Taylor, J. K. (1987). Quality assurance of chemical measurements. *Analytical Chemistry*, *53*(14), 1588A-1596A.

Taylor, S. C., Nadeau, K., Abbasi, M., Lachance, C., Nguyen, M., & Fenrich, J. (2019). The ultimate qPCR experiment: Producing publication quality, reproducible data the first time. *Trends in Biotechnology*, *37*(7), 761–774. https://doi.org/10.1016/j.tibtech.2018.12.002

US EPA. (2004). *Quality Assurance/Quality Control guidance for laboratories performing PCR analyses on environmental samples* (Data and Tools EPA 815-B-04-001). US EPA Office of Water. https://www.epa.gov/esam/quality-assurancequality-control-guidance-laboratories-performing-pcr-analyses-environmental

Verbyla, M. E., Symonds, E. M., Kafle, R. C., Cairns, M. R., Iriarte, M., Mercado Guzmán, A., Coronado, O., Breitbart, M., Ledo, C., & Mihelcic, J. R. (2016). Managing microbial risks from indirect wastewater reuse for irrigation in urbanizing watersheds. *Environmental Science & Technology*, *50*(13), 6803–6813. https://doi.org/10.1021/acs.est.5b05398

Westhaus, S., Weber, F.-A., Schiwy, S., Linnemann, V., Brinkmann, M., Widera, M., Greve, C., Janke, A., Hollert, H., Wintgens, T., & Ciesek, S. (2021). Detection of SARS-CoV-2 in raw and treated wastewater in Germany – suitability for COVID-19 surveillance and potential transmission risks. *Science of the Total Environment*, *751*, 141750. https://doi.org/10.1016/j.scitotenv.2020.141750

World Organisation for Animal Health (OIE). (2014). Chapter 2.2.5. Statistical approaches to validation. In *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals 2019* (pp. 210–221). World Organisation for Animal Health (OIE).

Ye, Y., Ellenberg, R. M., Graham, K. E., & Wigginton, K. R. (2016). Survivability, partitioning, and recovery of enveloped viruses in untreated municipal wastewater. *Environmental Science & Technology*, *50*(10), 5077–5085. https://doi.org/10.1021/acs.est.6b00876

Addendum: History of updates to the Protocol

Current Version: January 2022

- Clarification of concentration factor calculation and examples (Boxes 2.1, 4.1, 6.2)
- Clarification of repeatability acceptability thresholds
    - Based on an empirical analysis of standard curves through inter-laboratory method comparisons, the repeatability threshold of SD<0.5 Cq should be applied to dilutions greater than approximately 10 gene copies/reaction. A greater degree of variability is anticipated at lower concentrations and therefore would not be indicative of a standard curve that does not meet the repeatability criterion.
    - To hold replicates performed in duplicate to the same stringency of repeatability as triplicates (SD<0.5 Cq; n=3), the absolute delta Cq between replicates should be less than 0.8 Cq (|ΔCq|<0.8 Cq, n=2). More stringent laboratory-specific criteria can be prescribed through control charting.
- Fixed table and figure number references in Table of Contents and in text
- Minor grammatical edits throughout
- Addendum added to document changes made to the Protocol

Superseded Version: August 2021