
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Analytica Chimica Acta 1199 (2022) 339585

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

# Importance of sample input volume for accurate SARS-CoV-2 qPCR testing

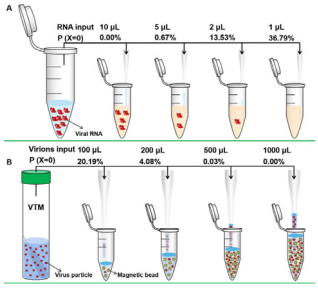
Yugan He [a, b], Tie Xie [a, b], Qihang Tu [b], Yigang Tong [a, b, *]

[a] Beijing Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology, Beijing, 100029, PR China
[b] College of Life Science and Technology, Beijing University of Chemical Technology, Beijing, 100029, PR China

## HIGHLIGHTS

- A Poisson distribution was used to illustrate the inconsistent performance of qPCR tests in detecting low viral load samples.
- False-negative qPCR results of clinical COVID-19 samples with a Ct $\geq$ 35 decreased by 50% after increasing the input of purified RNA from 2 to 10 μL.
- The consistency, accuracy, and robustness of nucleic acid testing for SARS-CoV-2 samples with low viral loads can be improved by increasing the sample input volume.

## GRAPHICAL ABSTRACT




## ARTICLE INFO

*Article history:*
Received 11 November 2021
Received in revised form
30 January 2022
Accepted 5 February 2022
Available online 7 February 2022

*Keywords:*
SARS-CoV-2
Nucleic acid testing
Sample input
Poisson distribution

## ABSTRACT

Nucleic acid testing is the most widely used detection method for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of the coronavirus disease 2019 (COVID-19) pandemic. Currently, a number of COVID-19 real-time quantitative reverse transcription PCR (qPCR) kits with high sensitivity and specificity are available for SARS-CoV-2 testing. However, these qPCR assays are not always reliable in detecting low viral load samples (Ct-value $\geq$ 35), resulting in inconclusive or false-negative results. Here, we used a Poisson distribution to illustrate the inconsistent performance of qPCR tests in detecting low viral load samples. From this, we concluded that the false-negative outcomes resulted from the random occurrences of sampling zero target molecules in a single test, and the probability to sample zero target molecules in one test decreased significantly with increasing purified RNA or initial sample input volume. At a given RNA concentration of 0.5 copy/μL, the probability of sampling zero RNA molecules decreased from 36.79% to close to 0.67% after increasing the RNA input volume from 2 to 10 μL. A SARS-CoV-2 qPCR assay with an LOD of 300 copies/mL was used to validate the improved consistency of the qPCR tests. We found that the false-negative qPCR results of clinical COVID-19 samples with a Ct $\geq$ 35 decreased by 50% after increasing the input of purified RNA from 2 to 10 μL. The consistency, accuracy, and robustness of nucleic acid testing for SARS-CoV-2 samples with low viral loads can be improved by increasing the sample input volume.



* Corresponding author. Beijing Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology, Beijing, 100029, PR China.
*E-mail address:* tongyigang@mail.buct.edu.cn (Y. Tong).

https://doi.org/10.1016/j.aca.2022.339585
0003-2670/

## 1. Introduction

The ongoing coronavirus disease 2019 (COVID-19) pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has resulted in more than 238 million infections worldwide and more than 4.8 million deaths as of October 16, 2021 [1,2]. Currently, SARS-CoV-2 continues to spread worldwide, including the emergence of several variants with increased transmissibility and disease severity, such as Gamma variant, Delta variant and Omicron variant [3–7]. In the absence of a sufficient supply of vaccines and effective therapy for the treatment of COVID-19 in some countries and regions, the test-trace-quarantine remains one of the most effective measures to restrain transmission [8,9]. After optimization by many laboratories, the real-time quantitative reverse transcription PCR (RT-qPCR or qPCR) assay has become the most widely adopted method for SARS-CoV-2 testing across the globe [10–12]. Several studies have focused on improving the sensitivity and specificity of the qPCR assay [11–13]; however, additional effort is needed to improve the diagnostic accuracy and consistency for SARS-CoV-2 detection, especially at low viral loads.

A number of commercially available SARS-CoV-2 qPCR kits have been developed and authorized by the National Medicine Products Administration (NMPA) for clinical diagnosis and the Food and Drug Administration (FDA) for emergency use. However, the performance of these assay kits in detecting low viral load samples (Ct-value $\geq$ 35) is inconsistent, resulting in inconclusive or false-negative results [14–18]. Reducing false-negative results is essential to effectively contain the spread of SARS-CoV-2. Here, we hypothesized that the inconsistent qPCR test results were due to randomness of sample pipetting, and the false-negative results were mainly due to sampling zero target molecules. This hypothesis is based on that 1) the number of occurrences of inputting *k* target DNA/RNA molecules into each reaction tube is random and independent, and that 2) test results are reported positive as long as each reaction tube contains $\geq$1 target molecules.

In statistics, a Poisson distribution is a probability distribution that is used to characterize how many times an independent event is likely to occur within some definite time or space [19–21]. Poisson distribution analysis has been used for absolute nucleic acid quantitation in digital PCR (dPCR) [22–25]. In low viral load samples, the probability of sampling zero target molecules in one test is relatively high; this can be estimated using a Poisson distribution analysis (Fig. 1). Here, we estimated the probability of sampling zero target molecules in a single qPCR test at various RNA concentrations across a wide range of RNA input volumes using a Poisson distribution. We also compared the performance of qPCR tests in detecting low SARS-CoV-2 load samples under various sample input volumes. We found that increasing either RNA or initial sample input volumes could significantly reduce the probability of sampling zero RNA molecules and thus improve the diagnostic accuracy and reliability of qPCR tests.

## 2. Materials and methods

### 2.1. SARS-CoV-2 RNA purification and concentration analysis

A SARS-CoV-2 pseudovirus particle, which was derived from a 2019-nCoV RT-qPCR Detection Kit (Fosun Pharma, Shanghai, China), was used as a simulated initial sample for purifying SARS-CoV-2 genomic materials. Two methods were used to purify SARS-CoV-2 RNA. 50 μL of SARS-CoV-2 RNA was extracted from 200 μL of SARS-CoV-2 pseudovirus particles using a QIAamp Viral RNA Mini Kit (Qiagen, Dusseldorf, Germany) following the manufacture's instructions. While for SARS-CoV-2 pseudovirus particle samples with a range of input volumes from 100 to 1000 μL, 20 μL of viral RNA was purified and concentrated using a Si–OH magnetic bead-based viral RNA extraction method, which was developed by our lab [26]. In comparison with silica column-based nucleic acid extraction method, magnetic beads-based nucleic acid extraction protocol is much simpler, faster, and does not require the use of specialized equipment [27–29]. More importantly, it is amenable to automated workflows for implementing high throughput RNA extraction [30–32].

The detailed procedure is as follows: (1) 20 μL of 10 mg/mL Si–OH magnetic beads (PuriMag, Xiamen, China) and 500 μL of lysis buffer were mixed in a 2.0 mL microfuge tube, followed by the addition of an indicated volume of initial sample; (2) the mixture was vortexed and incubated at around 25 °C for 10 min; (3) the sample tube was placed in a magnetic rack for 15 s, and then the supernatant was removed; (4) the magnetic beads remaining in the tube on the magnetic rack was washed with 500 μL of 80% ethanol two times and the supernatant was discarded; (5) the magnetic beads were then air dried for 5 min at around 25 °C; and (6) the viral RNA was eluted from the beads using 20 μL of nuclease-free water. The entire 20 μL eluate was then used as template for RT-qPCR testing.

### 2.2. RT-qPCR testing of the SARS-CoV-2 pseudovirus samples

Viral RNA was detected using a multiplex real-time RT-qPCR assay for the simultaneous detection of SARS-CoV-2-specific E and ORF1ab genes. To ensure test sensitivity and specificity, a 2019-nCoV RT-qPCR Detection Kit (20203400299), approved by the China NMPA, was used for the qPCR tests. The SARS-CoV-2 RNA sample, purified using a QIAamp Viral RNA Mini Kit, which had an approximate Ct value of 30.28 ± 0.24 corresponding to approximately 300 copies/μL, was first diluted 100-fold (3 copies/μL), 500-fold (0.6 copies/μL), and 1000-fold (0.3 copies/μL) in 1 × phosphate-buffered saline solution. The sample was then aliquoted for qPCR using different volumes. The qPCR reaction contained 7 μL of 2019-nCoV reaction buffer, 3 μL of enzyme mix, *x* μL of SARS-CoV-2 RNA, and 10-*x* μL of nuclease-free water, in a total volume of 20 μL. Ten replicates of each sample input volume were assayed.

For the SARS-CoV-2 RNA purified using magnetic beads, the qPCR reaction contained 7 μL of 2019-nCoV reaction buffer, 3 μL of enzyme mix and 20 μL of the purified RNA in a total volume of 30 μL. Five replicates of each sample input volume were assayed.

### 2.3. RT-qPCR testing of the SARS-CoV-2 clinical specimens

Clinical SARS-CoV-2 RNA samples were purified from 20 oropharyngeal swab specimens using a QIAamp Viral RNA Mini Kit from Wuhan Hospital, and stored at −80 °C. For the qPCR reaction, 2 μL or 10 μL of the viral RNA was added to a 20 μL reaction system, including 7 μL of 2019-nCoV reaction buffer, and 3 μL of enzyme mix, and 8 μL of nuclease-free water when the RNA input was 2 μL. The qPCR assay conditions included reverse transcription for 15 min at 50 °C, initial denaturation for 3 min at 95 °C, followed by 45 cycles of 5 s at 95 °C and 40 s at 60 °C. Fluorescence signals at this stage were collected using an ABI7500 qPCR machine (Applied Biosystems Inc).

### 2.4. Data analysis

The Poisson distribution data were generated using Microsoft Excel at the indicated $\lambda$ values; detailed information is provided in the Supplementary data. All figures were generated using either GraphPad Prism 8.3.0 or Origin 9.0. For *t*-test, $p < 0.05$ was considered significant.

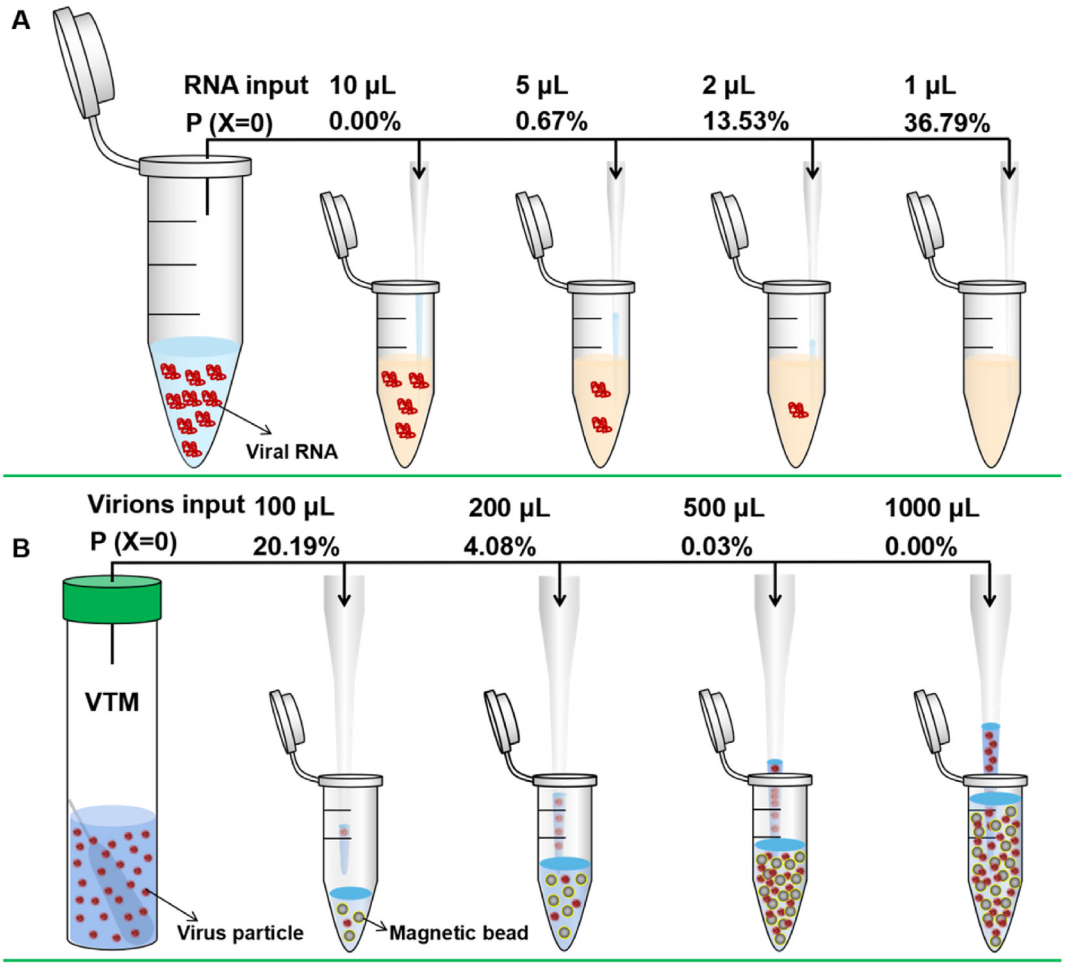



**Fig. 1.** Schematic representation of sampling across a range of sample input volumes. (A) Sampling viral RNA (1 copy/μL) to 0.2 mL qPCR reaction tubes across a range of input volumes for qPCR testing. (B) Sampling a swab sample containing virus particles (100 copies/mL) into 2.0 mL microtubes across a range of input volumes for subsequent viral RNA extraction. RNA bound to magnetic beads was eluted in 50 μL of nuclease-free water and 10 μL aliquots were then used in qPCR tests. The probabilities of sampling zero target molecules in each reaction tube ($P(X=0)$) were estimated using a Poisson distribution analysis at the indicated sample concentrations and input volumes.

## 3. Results and discussion

### 3.1. Increasing the input of purified viral RNA improves the accuracy of qPCR tests

In developing highly sensitive and specific qPCR assays for nucleic acids, sample input volume plays an important role in improving the test accuracy, particularly in detecting low viral load samples. When testing samples with a low viral load, qPCR tests can be inconsistent, resulting in false-negative results were reported [10–13]. This inconsistent performance can be illustrated using a Poisson distribution analysis. It is well accepted in digital PCR that the number of target molecules contained in each reaction unit (an aqueous droplet floating in oil or a fixed-location chamber) follows a Poisson distribution [33]. Similar to digital PCR, the number of nucleic acid copies input into each test (0.2 mL tube) in a qPCR assay also obeys a Poisson distribution, particularly for samples with low viral loads. The Poisson equations are presented in Equations (1)–(3) as follows:

$$\lambda = cV, \tag{1}$$

$$P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda}, (k=0,1,2\cdots), \tag{2}$$

$$\text{and for } k=0(\text{input zero RNA molecule}), P(X=0) = \frac{\lambda^0}{0!}e^{-\lambda} = e^{-\lambda}. \tag{3}$$

where, $P$ is the probability of sample $k$ RNA molecules in one test, $k$ is the copy number of the input RNA molecule in one test, $\lambda$ is the expected average copy number of the input RNA molecule, $c$ (copies/μL) is the viral RNA concentration, $V$ (μL) is the viral RNA input volume, and $e$ is Euler's constant ≈ 2.71828. In this study, we assumed that the qPCR assay is highly sensitive for viral RNA detection; that is, the viral RNA will test positive by qPCR as long as the input RNA molecules are ≥1 copy/test. In this case, for a SARS-CoV-2-infected individual to be identified positive by qPCR, we need to ensure that the quantity of the inputted viral RNA in one test is at least one copy. This would correspond to a probability of inputting zero RNA molecules in one test approaching 0.00% ($P(X=0) \approx 0.00\%$). Poisson distribution was used to estimate the probability of sampling zero viral RNA molecules in one test at indicated RNA concentrations across a wide range of RNA input

volumes. From the Poisson distribution analysis, we observed that the $P(X = 0)$ increases as the RNA concentration or RNA input volume decreases (Fig. 2).

As shown in Fig. 3A, as the viral RNA input volume increased, the probability of inputting zero RNA molecules in one test decreased significantly. According to the guidance from the GOV.UK, the desired limit of detection for COVID-19 nucleic acid testing is ≤ 100 SARS-CoV-2 virus particles per mL, corresponding to approximately 0.2 viral RNA copy/μL after RNA is purified [34]. In a SARS-CoV-2-containing sample with a known, low RNA concentration of 0.2 copy/μL, the copy number of the inputted RNA in each test at given input volumes can be estimated using a Poisson distribution. When 2 μL of the viral RNA was used for qPCR, the $P(X = 0)$ was ≥67.03%, indicating that the occurrence of false-negative results was ≥67.03%. With a viral RNA input of 5 μL, there was still a 36.79% chance of sampling zero RNA molecules; that is, only 63.21% of true positive results would test positive, which is lower than the required reproducibility of 95% in nucleic acid testing [35]. When the input volume was further increased to 10 μL, $P(X = 0)$ dropped to 13.53%. Only when the input volume was ≥20 μL, $P(X = 0)$ was ≤4.98%, corresponding to a 95% probability of a sample testing positive, which is acceptable in highly-sensitive nucleic acid tests. To ensure that $P(X = 0)$ approaches 0.00%, the required input volume would need to be ≥ 50 μL. However, in qPCR assays, the suggested total reaction volume is generally 20–50 μL. Therefore, by necessity, the RNA input volumes must be maintained at ≤ 40 μL. For more accurate and sensitive COVID-19 testing, an initial sample with a larger volume should be inputted, as discussed in detail in the following section.

Fig. 3B and C shows the probabilities of sampling different RNA copy numbers in one test under various input volumes at indicated RNA input concentrations of 1 and 0.1 copies/μL, respectively. The Poisson distribution analysis demonstrated that increasing RNA input volume significantly reduced the chance of sampling zero RNA molecules, and thus effectively reduced the inconclusive or false-negative results. More importantly, increasing viral RNA input volume from 2 to 10 μL at an indicated RNA concentration of 1 copy/μL resulted in an increase in $P(X \geq 3)$ from 32.33% to 99.72% (Fig. 3D). This observation exemplifies a marked increase in the chance of sampling more RNA molecules after increasing the sample input volume, which will contribute to a more accurate detection of target RNA.

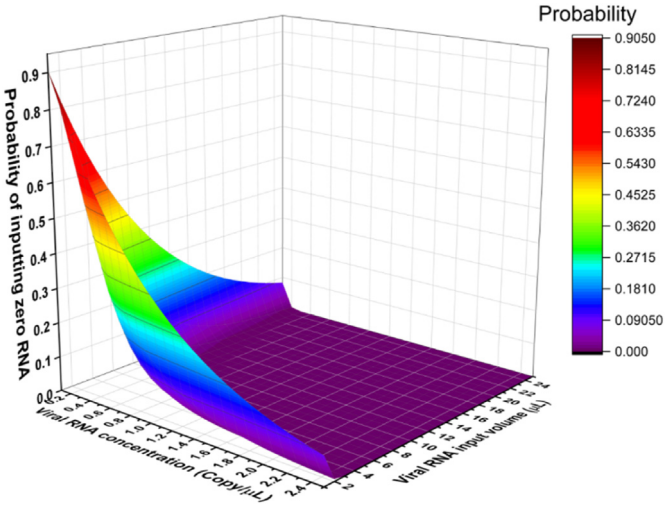



**Fig. 2.** Poisson distribution analysis to determine the probability of sampling zero viral RNA molecules in one reaction tube of a qPCR test at various RNA concentrations across a wide range of RNA input volumes. The viral RNA concentrations ranged from 0.1 to 2.5 copies/μL, and the RNA input volume ranged from 2 to 25 μL.

To validate the improved performance of nucleic acid tests after increasing viral RNA input volume, we performed qPCR using various RNA input volumes. In qPCR assay, SARS-CoV-2 RNA was detected by simultaneously targeting two different conserved segments of the viral ORF1ab gene and E gene, respectively [11,12]. Herein, SARS-CoV-2 RNA with a lower concentration of approximately 3 copies/μL, corresponding to a Ct value of ~35, was used as a template for the qPCR reaction. As the RNA input volume was increased from 2 μL to 10 μL, the number of replicates testing positive for SARS-CoV-2 E and ORF1ab genes increased from 8 (8/10) to 10 (10/10), and the Ct values decreased by 3 cycles (Ct of the E gene decreased from 38.67 to 35.64, and that of ORF1ab gene decreased from 37.49 to 34.47). Moreover, we found that the differences in Ct values among 1, 2, 5 and 10 μL input volumes were significant, even though the SARS-CoV-2 RNA input concentrations were the same. In addition, we found a greater dispersion of Ct values at smaller input volumes than at larger input volume (Fig. 4A). This finding is in accordance with published literature [10–12]. We believe that the inconsistent performance of the qPCR tests in detecting lower viral load samples is mainly caused by sampling different copy numbers of RNA molecules in each test, which could be well-illustrated by a Poisson distribution. Such a Poisson distribution analysis clearly showed that the $P(X \leq 2)$ in low viral load samples was very high at a small volume input, resulting in inconclusive or false-negative results.

We next compared the performance of qPCR tests using either 2 μL or 10 μL input volumes at various viral RNA concentrations. As shown in Fig. 4B and C, 100% (10/10) of the 100-fold (3 copies/μL) diluted sample yielded positive results with a 10 μL input; whereas 80% (8/10) of the tests yielded positive results with a 2 μL input. For samples diluted 500-fold (0.6 copies/μL), 70% (7/10) of the tests targeting the E gene and 80% (8/10) of the tests targeting the ORF1ab gene showed positive results with a 10 μL input. However, when 2 μL of the viral RNA was added, only 40% (4/10) of the tests were positive for both the E and ORF1ab genes. For samples diluted 1000-fold (0.3 copies/μL), the E and ORF1ab genes were still detected in 60% and 80% of the tests with a 10 μL input, respectively. However, decreasing the sample input volume to 2 μL resulted in 100% (10/10) of the tests being negative. Accordingly, we verified that increasing the sample input volume significantly improves the testing accuracy and consistency of qPCR assays, which is consistent with the results of the Poisson distribution.

### 3.2. Increasing the initial sample input improves the accuracy of qPCR tests

In addition to increasing the purified RNA input volume to improve qPCR test consistency and accuracy, we can also increase the initial sample input volume to enhance test reproducibility. We first estimated the concentration of the purified viral RNA (Equation (4)) under various initial sample input volumes at a given initial concentration, and then presented the probability of sampling exactly $k$ RNA molecules in one test by using the Poisson distribution analysis (Equation (5)).

$$\lambda = \frac{c_0 V_0 \beta}{V_1} V, \tag{4}$$

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, (k = 0, 1, 2 \cdots). \tag{5}$$

where $c_0$ (copies/μL) is the concentration (viral load) of the initial sample, $V_0$ (μL) is the initial sample input volume, $\beta$ (%) is the RNA

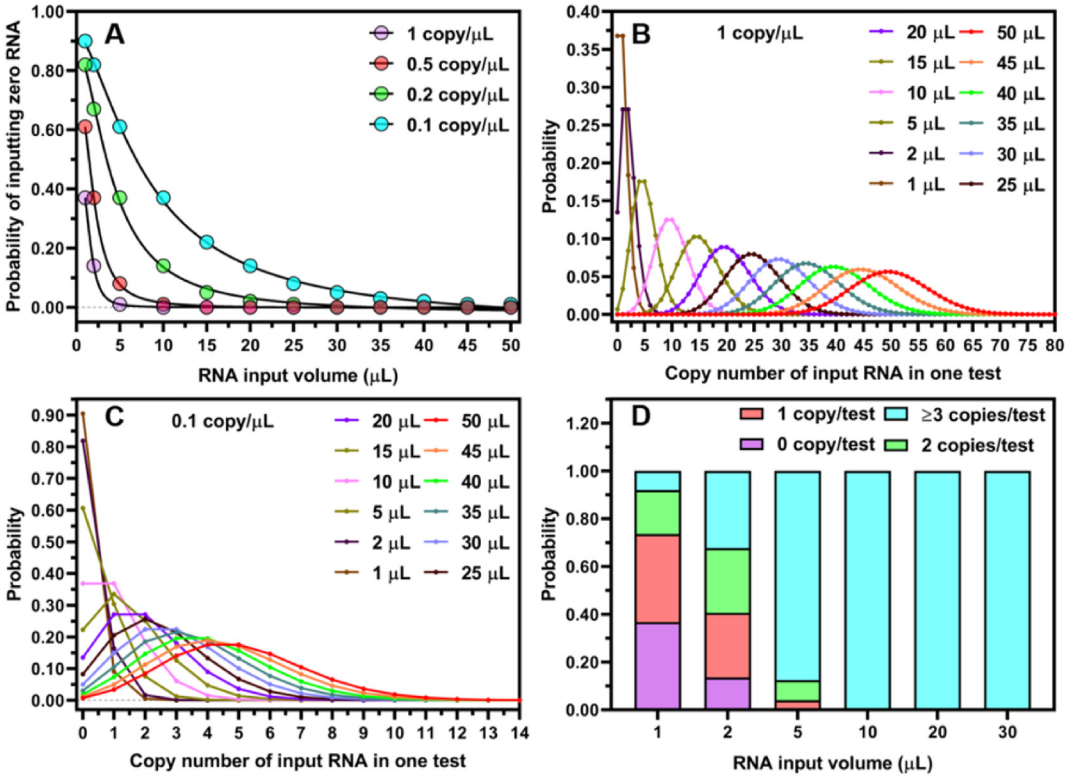



**Fig. 3.** Poisson distribution analysis of sampling viral RNA. (A) Probabilities of sampling zero RNA molecules in one test at different RNA concentrations with various RNA input volumes. (B) Probabilities of sampling a certain number of RNA molecules in one test under a range of viral RNA input volumes at an indicated RNA concentration of 1 copy/μL. (C) Probabilities of sampling a certain number of RNA molecules in one test under a range of viral RNA input volumes at an indicated RNA concentration of 0.1 copy/μL. (D) Probabilities of sampling 0, 1, 2, and ≥3 RNA molecules in one test under various RNA input volumes at a given RNA concentration of 1 copy/μL.

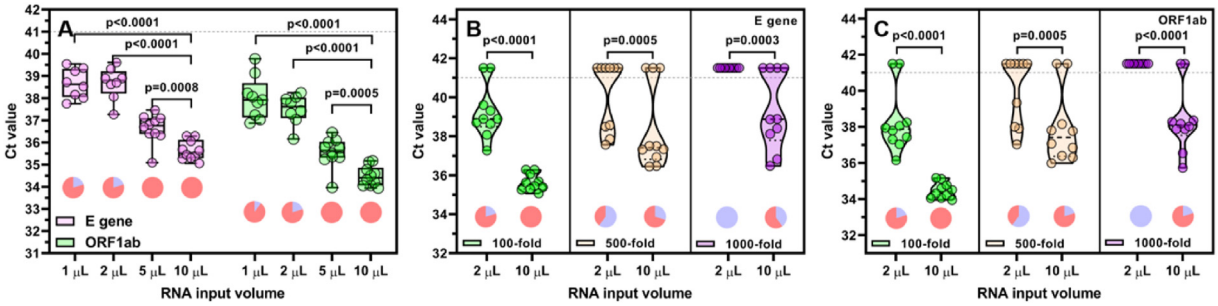



**Fig. 4.** qPCR performance with various viral RNA input volumes. (A) qPCR test results across a wide range of viral RNA (3 copies/μL) input volumes by using 10 replicates per input volume; (B–C) qPCR test results using either a 2 μL or 10 μL input volume with various RNA concentrations using 10 replicates per concentration. The purified SARS-CoV-2 RNA (300 copies/μL) was diluted with 1 × PBS 100-fold (3 copies/μL), 500-fold (0.6 copies/μL) and 1000-fold (0.3 copies/μL) as indicated. The red and blue circle wedges represent the proportion of samples testing positive and negative, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

extraction efficiency, and $V_1$ (μL) is the RNA elution volume. For various DNA/RNA extraction methods, the β is approximately 80%, and the suggested RNA elution volume $V_1$ is 50 μL [30]. If 50 μL of viral RNA is purified from an initial SARS-CoV-2 sample with a low viral load (down to 100 copies/mL = 0.1 copy/μL), and a 10 μL aliquot is used to perform qPCR, $\lambda$ and $P$ can be estimated as follows (Fig. 5A). When 100 μL of the initial sample was inputted, the $P(X = 0)$ was over 20.19%, indicating that the false-negative result was greater than 20%. To decrease the probability of sampling zero RNA molecules, we increased the initial sample input volume to 200 μL. Thus, $P(X = 0)$ was approximately 4.08%, which was less than the suggested value of 5.00% for the small probability event. This result indicates that 200 μL of the initial sample input is capable of meeting the requirement for qPCR to detect low viral load samples. To further eliminate the probability of sampling zero RNA molecules, 500 μL of the initial sample was used for RNA extraction. As a result, the $P(X = 0)$ significantly dropped to 0.03%, showing that $P$ decreased by 99.85% compared to that of the 100 μL-sample input. In addition, we also observed that the $P(X = 0)$ approached 0.00% when the initial sample input volumes were larger than 650 μL.

Based on the Poisson distribution, we can estimate the probabilities that a given test contains exactly $k$ RNA molecules, particularly for those containing zero target RNA molecules. As shown in Fig. 5A, at the indicated sample input volume, the probability of sampling $\lambda$ (round number) viral RNA molecules was the highest. $\lambda$ grew linearly as the initial sample input volume increased, which led to an exponential decrease in $P(X = \lambda)$. More importantly,

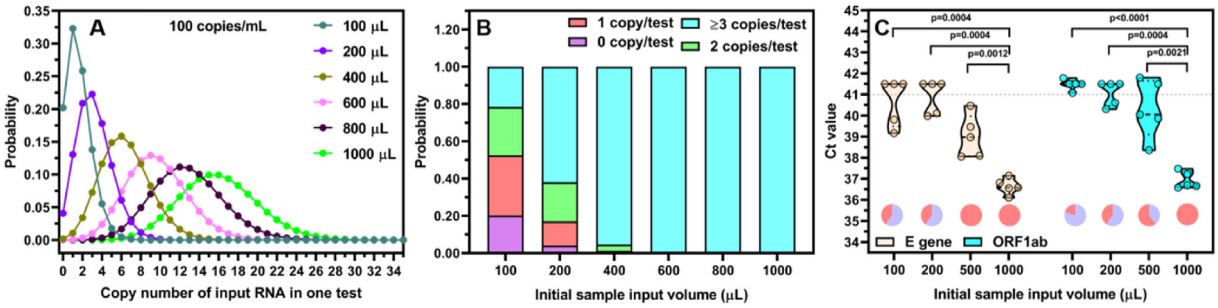



**Fig. 5.** Poisson distribution analysis of initial sample sampling. (A) Probabilities of sampling a certain number of RNA molecules in one test under a range of viral RNA input volumes at an indicated initial sample input concentration of 100 copies/mL. (B) Probabilities of sampling 0, 1, 2, and ≥3 RNA molecules in one test across a wide range of initial sample input volumes at a given concentration of 100 virus particles/mL. (C) Experimental validation of improved qPCR performance in detecting the SARS-CoV-2 E and ORF1ab genes in low viral load samples (0.3 copies/μL) after increasing the initial sample input volumes. Viral RNA was isolated and concentrated from an indicated volume of SARS-CoV-2 pseudovirus sample, followed by elution with 20 μL nuclease-free water for downstream qPCR testing. The red and blue circle wedges represent the proportion of samples testing positive and negative, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

increasing the initial sample input volume reduced the occurrence of sampling zero RNA molecules, which would enable diagnostic accuracy and effectively avoid false-negative results. For instance, in a swab sample with a low viral load of 100 copies/mL, the $P(X = 0)$ dropped from 20.19% to 0.17% as the sample input volume increased from 100 μL to 400 μL, and approached 0.00% when 800 μL of the initial sample was loaded. On the other hand, we observed that increasing the initial sample input volume led to a much higher probability of inputting more target molecules; for example, the $P(X \geq 3)$ increased from 21.66% to 99.97% after the initial sample input volume increased from 100 μL to 800 μL (Fig. 5B). However, we also found a diminishing returns effect, which shows that there is a point where an increased level of inputs does not equal to an equal increase level of outputs [36,37]. At this point, increasing sample input volume will no longer improve test accuracy and consistency. As shown in Figs. 3D and 5B, when the RNA (1 copy/μL) input volume was >10 μL or the initial sample (100 virus particles/mL) input volume was >600 μL, the $P(X \geq 3)$ has approached 100% and hardly increase. Loading a larger volume of initial sample for RNA extraction can effectively concentrate the target molecules and thus reduce the probability of zero sample target molecules.

Moreover, we investigated the performance of qPCR tests across a wide range of initial sample input volumes using a low viral load of a SARS-CoV-2 pseudovirus. As shown in Fig. 5C, the qPCR test results demonstrated that increasing the initial sample input volume from 100 μL to 1000 μL increased the test reliability in detecting low viral load samples. When the initial sample input volume was 100 μL, the E gene tested positive in 40% (2/5) of the replicates, while the ORF1ab gene tested positive in only 20% (1/5) of the replicates. An input of 500 μL of the initial sample significantly decreased the inconclusive results, resulting in 100% (5/5) and 60% (3/5) positive results for detecting the E and ORF1ab genes, respectively. While 1000 μL of the initial sample was used to isolate and concentrate viral RNA, 100% (5/5) of the tests showed positive results for both the E gene (36.64 ± 0.39 Ct) and ORF1ab gene (36.92 ± 0.40 Ct), demonstrating a reliable performance in detecting low viral load samples. Accordingly, we demonstrated that increasing the initial sample input volume significantly increases the chance of sampling more RNA molecules, and thus greatly improves the analytical accuracy of nucleic acid-based tests.

### *3.3. Poisson distribution analysis of sampling the initial sample without RNA extraction*

Owing to the significant savings in time and cost, the RNA-extraction-free qPCR assays have been developed and are being used by some research laboratories [38–40]. However, it is a great challenge for researchers to establish an RNA-extraction-free qPCR assay with high sensitivity for COVID-19 testing, owing to the high probability of sampling zero RNA molecules in low viral load samples with a small-volume input. As estimated by Poisson distribution analysis, at an indicated concentration of 100 copies/mL, the $P(X = 0)$ is 60.65% with a 5 μL input, dropping to 13.53% after the input volume was increased to 20 μL. For samples with a viral load of 500 copies/mL, the $P(X = 0)$ was >8.21% with a 5 μL input, and it approached 0.00% with a 20 μL input. When the initial concentration was 1000 copies/mL, a 10 μL input would enable a >99.99% chance of sampling ≥1 virus particles and reporting a positive test result (Fig. S1). In short, we demonstrated that the standard qPCR protocol with RNA-extraction is still a preferred, highly sensitive test for COVID-19 diagnosis, and the RNA-extraction-free RT-qPCR assay is more suitable for detection of samples with viral loads >1000 copies/mL.

### *3.4. Clinical validation of improved qPCR test performance after increasing the RNA input volume*

We further investigated the performance of qPCR tests with an input volume of 2 μL and 10 μL RNA using 20 known SARS-CoV-2 RNA clinical samples, including 11 positive and 9 negative samples (Fig. S2). For the four samples with Ct values ≤ 35, both SARS-CoV-2 genes tested positive with either a 2 μL input (E gene: 4/4, 29.11–36.29 Ct; ORF1ab: 4/4, 28.74–36.45 Ct) or a 10 μL input (E gene: 4/4, 26.83–34.08 Ct; ORF1ab: 4/4, 26.45–33.91 Ct), suggesting that sample input volume has no effect on qPCR results when detecting higher viral load samples (≥3 viral RNA copies/μL). For the seven samples with Ct values > 35, the E gene tested 100% positive results (7/7, 38.19–39.17 Ct) with a 10 μL input, and 43% positive (3/7, 38.71–38.73 Ct) with a 2 μL input. In turn, the ORF1ab gene tested 86% positive (6/7, 37.53–39.68 Ct) with a 10 μL input, and 57% positive (4/7, 38.13–39.32 Ct) with a 2 μL input. The nine negative samples with no SARS-CoV-2 RNA, were all tested negative, demonstrating 100% specificity for COVID-19 testing. We again showed that increasing the sample input volume contributes to a more accurate and highly sensitive detection of SARS-CoV-2 RNA, particularly in samples with low viral loads.

## 4. Conclusions

Overall, we believe that the inconsistency of qPCR tests in detecting low viral load samples (Ct ≥ 35) is caused by the

occurrence of sampling zero target molecules, which results in false-negative results. We estimated the probability of sampling different numbers of target molecules in a single qPCR test using a Poisson distribution analysis.The results showed that the performance of qPCR in detecting low viral load samples can be significantly improved by reducing the probability of sampling zero target molecules by increasing the RNA input volume, particularly by increasing the initial sample input volume. We suggest using $\geq$10 μL of purified viral RNA for qPCR tests and 500−1000 μL of the initial sample for viral RNA extraction. Moreover, we suggest using the gold-standard RNA-extraction qPCR test for highly sensitive COVID-19 diagnosis over that of RNA-extraction-free qPCR.

## CRediT authorship contribution statement

**Yugan He:** performed experiments, analyzed data and prepared figures, participated in sample handling and organization, conceived the study and wrote the manuscript. **Tie Xie:** performed experiments, participated in sample handling and organization. **Qihang Tu:** performed experiments, participated in sample handling and organization. **Yigang Tong:** supervised the study, All authors participated in manuscript editing and approved the manuscript.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This research was supported and funded by the National Key Research and Development Program of China (2018YFA0903000) and China MoST Emergency Project on COVID-19 (2020YFC084800).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2022.339585.

## References

[1] https://www.who.int/.
[2] F. Wu, S. Zhao, B. Yu, et al., A new coronavirus associated with human respiratory disease in China, Nature 579 (2020) 265−269.
[3] https://www.cdc.gov/.
[4] David Adam, What scientists know about new, fast-spreading coronavirus variants, Nature 594 (2021) 19−20.
[5] F.G. Naveca, V. Nascimento, V.C. de Souza, et al., COVID-19 in Amazonas, Brazil, was driven by the persistence of endemic lineages and P.1 emergence, Nat. Med. 27 (2021) 1230−1238.
[6] X.T. Zhang, S.J. Wu, B.L. Wu, et al., SARS-CoV-2 Omicron strain exhibits potent capabilities for immune evasion and viral entrance, Signal Transduct Target Ther 6 (2021) 430.
[7] H. Ledford, How severe are Omicron infections? Nature 600 (2021) 577−578.
[8] L. Mutesa, P. Ndishimye, Y. Butera, et al., A pooled testing strategy for identifying SARS-CoV-2 at low prevalence, Nature 589 (2021) 276−280.
[9] S. Contreras, J. Dehning, M. Loidolt, et al., The challenges of containing SARS-CoV-2 via test-trace-and-isolate, Nat. Commun. 12 (2021) 1−13.
[10] C.B.F. Vogels, A.F. Brito, A.L. Wyllie, et al., Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT-qPCR primer-probe sets, Nat Microbiol 5 (2020) 1299−1305.
[11] E. Kudo, B. Israelow, C.B.F. Vogels, et al., Detection of SARS-CoV-2 RNA by multiplex RT-qPCR, PLoS Biol. 18 (2020), e3000867.
[12] Y.J. Jung, G.S. Park, J.H. Moon, et al., Comparative analysis of primer−probe sets for RT-qPCR of COVID-19 causative virus (SARS-CoV-2), ACS Infect. Dis. 6 (2020) 2513−2523.
[13] C.B.F. Vogels, M.I. Breban, I.M. Ott, et al., Multiplex qPCR discriminates variants of concern to enhance global surveillance of SARS-CoV-2, PLoS Biol. 19 (2021), e3001236.
[14] W.B. Blake, S.H. JessiCa, G.G. Cameron, et al., Distribution of SARS-CoV-2 PCR cycle threshold values provide practical insight into overall and target-specific sensitivity among symptomatic patients, Am. J. Clin. Pathol. 154 (2020) 479−485.
[15] X.M. Hu, L.X. Zhu, Y.J. Luo, et al., Evaluation of the clinical performance of single-, dual-, and triple-target SARS-CoV-2 RT-qPCR methods, Clin. Chim. Acta 511 (2020) 143−148.
[16] L. Yu, S.S. Wu, X.W. Hao, et al., Limits of detection of 6 approved RT-PCR kits for the novel SARS-Coronavirus-2 (SARS-CoV-2), Clin. Chem. 66 (2020) 977−978.
[17] S.O. Nuno, C.N. Margarida, Implication of SARS-CoV-2 evolution in the sensitivity of RT-qPCR diagnostic assays, Lancet Infect. Dis. 21 (2021) 166−167.
[18] R. Rezgar, M. Nariman, A. Mohammad, rRT-PCR for SARS-CoV-2: analytical considerations, Clin. Chim. Acta 516 (2021) 1−7.
[19] G.S. Yen, B.S. Fujimoto, T. Schneider, et al., Statistical analysis of nonuniform volume distributions for droplet based digital PCR assays, J. Am. Chem. Soc. 141 (2019) 1515−1525.
[20] R.B. Dorazio, M.E. Hunter, Statistical models for the analysis and design of digital polymerase chain reaction (dPCR) experiments, Anal. Chem. 87 (2015) 10886−10893.
[21] C. Hindson, J. Chevillet, H. Briggs, et al., Absolute quantification by droplet digital PCR versus analog real-time PCR, Nat. Methods 10 (2013) 1003−1005.
[22] F.T. Yu, L.T. Yan, N. Wang, et al., Quantitative detection and viral load analysis of SARS-CoV-2 in infected patients, Clin. Infect. Dis. 71 (2020) 793−798.
[23] X. Cui, L. Wu, Y. Wu, et al., Fast and robust sample self-digitization for digital PCR, Anal. Chim. Acta 1107 (2020) 120−134.
[24] Y.Y. Jiang, A. Manz, W.M. Wu, Fully automatic integrated continuous-flow digital PCR device for absolute DNA quantification, Anal. Chim. Acta 1125 (2020) 50−56.
[25] Y.F. Ning, X. Cui, C. Yang, et al., A self-digitization chip integrated with hydration layer for low-cost and robust digital PCR, Anal. Chim. Acta 1055 (2019) 65−73.
[26] Y.G. He, T. Xie, Y.G. Tong, Rapid and highly sensitive one-tube colorimetric RT−LAMP assay for visual detection of SARS-CoV-2 RNA, Biosens. Bioelectron. 187 (2021) 113330.
[27] Haile, S., Nikiforuk, A.M., Pandoh, P.K., et al. Optimization of magnetic bead-based nucleic acid extraction for SARS-CoV-2 testing using readily available reagents. J. Virol. Methods 299, 114339.
[28] H. Yamaguchi, K. Matsuda, M. Uehara, et al., A novel automated device for rapid nucleic acid extraction utilizing a zigzag motion of magnetic silica beads, Anal. Chim. Acta 906 (2016) 1−6.
[29] R.B. Ami, A. Klochendler, M. Seidal, et al., Large-scale implementation of pooled RNA extraction and RT-PCR for SARS-CoV-2 detection, Clin. Microbiol. Infect. 26 (2020) 1248−1253.
[30] B. Netta, B.A. Roni, S. Tal, et al., Lessons from applied large-scale pooling of 133816 SARS-CoV-2 RT-PCR tests, Sci. Transl. Med. 13 (2021), eabf2823.
[31] N. Shental, S. Levy, V. Wuvshet, et al., Efficient high-throughput SARS-CoV-2 testing to detect asymptomatic carriers, Sci. Adv. 6 (2020), eabc5961.
[32] K.H. Chan, W.C. Yam, C.M. Pang, et al., Comparison of the NucliSens easyMAG and Qiagen BioRobot 9604 nucleic acid extraction systems for detection of RNA and DNA respiratory viruses in nasopharyngeal aspirate samples, J. Clin. Microbiol. 46 (2008) 2195−2199.
[33] Y. Zhang, H. Noji, Digital bioassays: theory, applications, and perspectives, Anal. Chem. 89 (2017) 92−101.
[34] https://www.gov.uk/.
[35] A. Ceci, C. Muñoz-Ballester, A.N. Tegge, et al., Development and implementation of a scalable and versatile test for COVID-19 diagnostics in rural communities, Nat. Commun. 12 (2021) 4400−4412.
[36] R. Lei, K.X. Ye, Z.L. Gu, et al., Diminishing returns in next-generation sequencing (NGS) transcriptome data, Gene 557 (2015) 82−87.
[37] H.A. Johnson, Diminishing returns on the road to diagnostic certainty, JAMA 265 (1991) 2229−2231.
[38] I. Smyrlaki, M. Ekman, A. Lentini, et al., Massive and rapid COVID−19 testing is feasible by extraction-free SARS-CoV-2 RT-PCR, Nat. Commun. 11 (2020) 4812.
[39] N. Merindol, G. Pépin, C. Marchand, et al., SARS-CoV-2 detection by direct rRT-PCR without RNA extraction, J. Clin. Virol. 128 (2020) 104423.
[40] S.A. Byrnes, R. Gallagher, A. Steadman, et al., Multiplexed and extraction-free amplification for simplified SARS-CoV-2 RT-PCR tests, Anal. Chem. 93 (2021) 4160−4165.