Clinical Performance of Self-Collected Purified Water Gargle for Detection of Influenza A Virus Infection by Real-Time RT-PCR

Guiling Li1,3,*, Tianyang Tan3,5,*, Luting Chen3,5, Jiaqi Bao3,5, Dongsheng Han3,5, Fei Yu4,5

1Department of Clinical Laboratory, Sir Run Run Shaw Hospital of Medicine Zhejiang University, Hangzhou, 310003, People’s Republic of China; 2Key Laboratory of Precision Medicine in Diagnosis and Monitoring Research of Zhejiang Province, Hangzhou, 310003, People’s Republic of China; 3Institute of Laboratory Medicine, Zhejiang University, Hangzhou, 310003, People’s Republic of China; 4Department of Laboratory Medicine, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, 310003, People’s Republic of China; 5Zhejiang Key Laboratory of Clinical in vitro Diagnostic Techniques, Hangzhou, 310003, People’s Republic of China

*These authors contributed equally to this work

Correspondence: Dongsheng Han; Fei Yu, Department of Laboratory Medicine, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, 310003, People’s Republic of China, Email 1321269@zju.edu.cn; yufeizy@zju.edu.cn

Purpose: Self-collected specimens are increasingly being used as alternatives to swab-based methods for the detection of respiratory viruses. While saliva is well accepted, gargle specimens are a potential alternative with characteristics that are more favorable for laboratory handling. This study assessed the performance of gargle specimens in the detection of influenza A viruses (IAVs).

Patients and Methods: We performed a prospective head-to-head comparison between combined nasopharyngeal and oropharyngeal swabs (NPS&OPS) and purified water gargle (PWG) among adult outpatients with febrile respiratory symptoms to detect IAVs using real-time RT-PCR during two influenza seasons.

Results: During study periods 1 (July 13 to 26, 2022, H3N2 predominated) and 2 (February 25 to March 10, 2023, H1N1 pdm09 predominated), a total of 459 patients were recruited. The overall agreement between the NPS&OPS and PWG was 85.0% (390/459, \(\kappa = 0.697\)), with 88.0% in period 1 and 82.6% in period 2. The detection rate of IAVs in PWG (51.6%, 237/459) was lower than that in NPS&OPS (62.3%, 286/459) \((p < 0.0001)\). The overall sensitivity and specificity were 96.6% (93.7–98.3%) and 100% (97.1–100%) in NPS&OPS and were 80.1% (75.0–84.4%) and 100% (97.1–100%) in PWG, respectively. Among the 227 pairs of concordant positive specimens, cycle threshold (Ct) values were significantly lower in NPS&OPS than in PWG (median Ct values: 24.2, 28.2, \(p < 0.0001\)).

Conclusion: Although self-collected PWG specimens offer acceptable performance for IAVs molecular testing, NPS&OPS remain a reliable option. Given the convenience of collection, nonviscous gargles are recommended for viral detection during emergencies or under specific conditions.

Keywords: gargle, nasopharyngeal swab, oropharyngeal swab, detection, influenza A virus, rRT-PCR

Introduction

Collecting respiratory tract specimens in time to detect respiratory viruses can provide valuable information, helping to manage patients properly and prevent transmission. However, the recommended method for specimen collection is debatable. The selection of a sampling method must balance epidemiological sensitivity against the feasibility, costs, time required for specimen collection, and other aspects.1 For example, the “gold standard” sample types for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) detection defaulted for early in the pandemic to nasopharyngeal swab (NPS) or oropharyngeal swab (OPS), as these are established diagnostic practices used in identifying other respiratory viruses and are often recommended by clinical guidelines.2–5 However, these specimens can be costly, labor-intensive, invasive, and put health personnel at risk by close contact during sample collection,4 leading to the consideration of additional self-collected and less-invasive specimen types. Saliva has
been proposed as an acceptable specimen type due to some advantages such as easy to self-collect, more acceptable to patients, less resource intensive and low transmission risk.\textsuperscript{4–7}

We evaluated the use of saliva for detecting SARS-CoV-2 and common respiratory pathogens,\textsuperscript{8,9} and our findings revealed that saliva could be easily provided by patients and was well accepted. However, saliva still has shortcomings in practical applications; for instance, some patients who experience dry mouth cannot spit out a sufficient volume of saliva and the obtained saliva with different viscosities needs to be liquefied and centrifuged during pretreatment, which is both time consuming and increases the detection difficulty for home detection and automatic equipment detection.\textsuperscript{10–15} In view of this, we speculated that gargle may be a potential alternative that is easier to obtain from individuals having difficulty in providing saliva and more favorable for laboratory handling. Gargle samples appear more attractive because they are non-viscous in nature, minimizing cross-contamination during resuspension and transfer, which poses the risk of false-positive results and reduces the chance of clogging liquid-handling systems that would result in testing failure. According to previous studies, gargles have been used to detect respiratory viruses, especially SARS-CoV-2,\textsuperscript{5,10–18} while their application in other respiratory virus testing is more limited.\textsuperscript{19,20} In this study, we performed a prospective head-to-head comparison of the clinical performance of self-collected gargle specimens in detecting influenza A Virus (IAVs) by real-time reverse transcription PCR (rRT-PCR) during the influenza season.

**Materials and Methods**

**Subjects**

This prospective diagnostic validity study was conducted in (1) period 1: flu caused by H3N2 virus predominated from July 13 to 26, 2022; (2) period 2: flu caused by H1N1pdm09 virus predominated from February 25 to March 10, 2023. Patients who visited the Department of Fever Clinics, First Affiliated Hospital, Zhejiang University School of Medicine, were included in this study. The inclusion criteria were at least one of the following focal or systemic symptoms: fever, cough, rhinorrhea or rhinoyon, headache, muscle soreness, and feeling unwell, which were considered to be screened for IAVs. The exclusion criteria were inability to provide qualified specimens, refusal to sample, and inability to provide written informed consent. Demographic information, symptoms and signs, date of symptom onset, and underlying diseases were obtained from clinical records. All the data were reviewed by a trained team of physicians. This study was conducted in accordance with the principles of the Declaration of Helsinki and approved by the Clinical Research Ethics Committee of the First Affiliated Hospital of the Zhejiang University School of Medicine.

**Specimen Collection**

Before sampling, it was confirmed that the patient did not drink, eat, gargle, or have other similar behaviors within half an hour, which might have affected the sampling quality. Combined NPS and OPS (NPS&OPS) specimens were collected by skilled medical personnel using a standardized method. NPS were collected first, followed by OPS, and finally gargled at an interval of 5 min (Figure 1). A thin swab with a flocked rayon fiber tip (Yocon China) was inserted into one nostril until it reached the back of the nasopharyngeal cavity and rotated before removal, while the OPS was directed toward the rear wall of the oropharynx and the left and right tonsillar arches with a rayon throat swab (Yocon China); the swab was wiped three times before removal. Immediately after collection, the swabs were inserted into the same tube containing 3 mL viral transport medium (Yocon China).\textsuperscript{21} The purified water gargle (PWG) specimens were collected by the participants themselves. The personnel observed and showed the participants how to collect the gargles and reminded them not to swallow the mouthwash. Each participant received a sterile container containing 5 mL of Wahaha\textsuperscript{60} drinking purified water and was instructed to rinse their mouth for 5 s, tilt their head back and gargle for 5 s, repeat this cycle once, and then spit out the entire contents back into the same container. The specimens were stored at 4 °C and processed within 8 h of collection.

**Laboratory Testing**

The specimens were vortexed ten times using a vortexer for 10 times. Then, 200 μL of the specimen was taken after vortex mixing, as well as negative and positive controls. Viral RNAs were extracted simultaneously on the sampling day using an
Automated Nucleic Acid Extraction System (EX3600, Zhijiang Bio-Tech, Shanghai, China) based on the paramagnetic bead method. rRT-PCR was performed using a China Food and Drug Administration (CFDA) approved commercial kit specific for IAV detection with ABI 7500 Real-Time PCR instrument (Foster City, CA, USA). Positive nucleic acid specimens were tested again using the H3N2 and H1N1 pdm09 subtyping kits. The extraction kit (Cat. No. ZME0044) and testing kits (Cat. No.: WRR005102, WRR042004, WRR013902) were purchased from Zhijiang Biotechnology Co., Ltd. (Shanghai, China). Operative and result assessments were performed in accordance with the manufacturer’s instructions. Sample volumes were 200 µL, with elution volumes of 50 µL, and 5µL used for PCR, in a total reaction volume of 25 µL. The detection limit of the assay was approximately 1000 copies/mL. All the amplification curves were manually inspected. Cycle threshold

**Figure 1** Recruitment flow chart and summary of NPS&OPS and PWG specimens results of IAVs detection by rRT-PCR in 459 patients.

**Abbreviations:** VTM, viral transport medium; IAVs, influenza A virus; NPS&OPS, combined nasopharyngeal and oropharyngeal swab; PWG, purified water gargle.
(Ct) values were used as a surrogate for the viral load. Specimens with Ct values ≤38.0 were considered positive, and >38.0 were considered negative.

**Statistical Analysis**
Data were described as percentages for categorical variables and medians with interquartile ranges (IQR) for continuous variables. Agreement between PWG and NPS&OPS was performed using κ statistics. The detection rates of paired specimens were compared using McNemar’s test. For sensitivity and specificity, the reference standard for patients were considered infected with IAVs if any specimen tested positive. Online Vassar Stats website (http://vassarstats.net/index.html) was used to calculate detection rate and their respective confidence intervals. Ct values were compared using the Wilcoxon matched-paired signed-rank test. The correlation of Ct values between PWG and NPS&OPS was assessed using the Spearman correlation coefficient. Statistical significance was set at p <0.05. All statistical analyses were performed using GraphPad PRISM® version 9.5.1 or SPSS 18.0.

**Results**

**Patient Characteristics**
During the two study periods, 459 patients were enrolled, comprising 200 patients in period 1 and 259 patients in period 2. Patient characteristics are shown in Table S1. The median age of the patients was 33 years, and 48.8% (224/459) were female. A total of 352 (76.7%) patients visited our hospital and were sampled within 48 hours of symptom onset. The most common symptom was fever (95.0%, 436/459), followed by cough (55.3%, 254/459). A total of 152 patients (33.1%) had underlying diseases such as hypertension and diabetes mellitus.

**Analysis of the Agreement Between NPS&OPS and PWG**
All 459 eligible patients showed valid results for the paired specimens using rRT-PCR. The overall agreement between NPS&OPS and PWG was 85.0% (390/459, κ 0.697, 95% CI 0.633–0.761), with 88.0% (176/200, κ 0.762, 95% CI 0.675–0.849) in period 1 and 82.6% (214/259, κ 0.624, 95% CI 0.527–0.721) in period 2, respectively (Figure 1). The detection rate of IAVs in PWG (51.6%, 237/459) was significantly lower than that in NPS&OPS (62.3%, 286/459) (p < 0.0001), with similar results (52.5% (105/200) vs 41.5% (83/200); p < 0.0001) in period 1 and results (59.5% (154/259) vs 69.9% (181/259); p < 0.0001) in period 2.

**Comparison of Sensitivity and Specificity**
In period 1, 106 (53.0%, 106/200) patients were considered infected with IAVs (100% H3N2 by subtyping). The sensitivities of NPS&OPS and PWG were 99.1% (95% CI, 94.1–100%) and 78.3% (95% CI 69.0–85.5%), respectively. In period 2, 190 (73.3%, 190/259) patients were considered infected with IAVs (89.5% H1N1 pdm09 and 10.5% H3N2 by subtyping). The sensitivities of the NPS&OPS and PWG were 95.3% (95% CI, 90.9–97.7%) and 81.1% (95% CI, 74.6–86.2%), respectively. The overall sensitivity and specificity of NPS&OPS were 96.6% (95% CI, 93.7–98.3%) and 100% (95% CI, 97.1–100%), respectively. The overall sensitivity and specificity of PWG were 80.1% (95% CI 75.0–84.4%) and 100% (95% CI, 97.1–100%), respectively (Table 1). We examined the association between age, gender, the onset of symptoms, the presence of the symptoms and the sensitivities of specimen types (Table S2). For PWG, there was a significant association between gender and sensitivities (male 86.1% vs female 73.8%; p = 0.008). No significant associations were observed for other characteristics.

**Viral Load Analysis**
Higher Ct values were observed in the majority of PWG specimens (Figure 2A). Among the 227 patients with concordant positive pairs, Ct values were significantly higher in PWG than in NPS&OPS (median Ct values, 28.2 versus 24.2, Δ=4.0, p < 0.0001) (Figure 2B). Notably, 26.4% (60/227) of the patients had lower Ct values in their PWG than in the NPS&OPS. A weak correlation of IAVs Ct values was found between paired specimens (r 0.188, p = 0.005) (Figure 2C). When patients had lower Ct value intervals in their NPS&OPS, the difference between paired specimens was greater
Except 10 infected patients only had positive PWG results, among the 286 infected patients positive in NPS&OPS, the positive rates of PWG were 92.3% (48/52), 83.7% (82/98), 73.3% (66/90) and 67.4% (31/46) in different NPS&OPS Ct value groups (<20, 20–25, 25–30 or ≥30), respectively (Figure 2E).}

Table 1 Results of IAVs Detection by rRT-PCR in NPS&OPS and PWG Specimens When Compared with Patients’ Infection Status

<table>
<thead>
<tr>
<th>Patient Infected with IAVs</th>
<th>Period 1</th>
<th></th>
<th>Period 2</th>
<th></th>
<th>Overall</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>YES</td>
<td>NO</td>
<td>Total</td>
<td>YES</td>
<td>NO</td>
<td>Total</td>
</tr>
<tr>
<td>Any specimen tested</td>
<td>106</td>
<td>94</td>
<td>200</td>
<td>190</td>
<td>69</td>
<td>259</td>
</tr>
<tr>
<td>NPS&amp;OPS Positive</td>
<td>105</td>
<td>0</td>
<td>105</td>
<td>181</td>
<td>0</td>
<td>181</td>
</tr>
<tr>
<td>Negative</td>
<td>1</td>
<td>94</td>
<td>95</td>
<td>9</td>
<td>69</td>
<td>78</td>
</tr>
<tr>
<td>Sen. (95% CI) (%)</td>
<td>99.1 (94.1–100)</td>
<td>95.3 (90.9–97.7)</td>
<td>96.6 (93.7–98.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spe. (95% CI) (%)</td>
<td>100 (95.1–100)</td>
<td>100 (93.4–100)</td>
<td>100 (97.1–100)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PWG Positive</td>
<td>83</td>
<td>0</td>
<td>83</td>
<td>154</td>
<td>0</td>
<td>154</td>
</tr>
<tr>
<td>Negative</td>
<td>23</td>
<td>94</td>
<td>117</td>
<td>36</td>
<td>69</td>
<td>105</td>
</tr>
<tr>
<td>Sen. (95% CI) (%)</td>
<td>78.3 (69.0–85.5)</td>
<td>81.1 (74.6–86.2)</td>
<td>80.1 (75.0–84.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spe. (95% CI) (%)</td>
<td>100 (95.1–100)</td>
<td>100 (93.4–100)</td>
<td>100 (97.1–100)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: IAVs, influenza A virus; NPS&OPS, combined nasopharyngeal and oropharyngeal swab; PWG, purified water gargle; Sen., sensitivity; Spe., specificity; 95% CI, 95% confidence interval.

Figure 2 Ct values of IAVs detection by rRT-PCR in NPS&OPS and PWG specimens. (A) Comparison of Ct values of paired specimens. (B) Comparison of Ct values of positive specimens. (C) Correlation of Ct values between PWG and NPS&OPS. (D) Comparison of Ct values between paired specimens by different NPS&OPS Ct values intervals. (E) Positive rates of PWG specimens in different NPS&OPS Ct values intervals. Only patients with concordant positive sample pairs were included in B and C. It should be noted that the higher the Ct value, the lower the viral load.

Abbreviations: Ct, cycle threshold; IAVs, influenza A virus; NPS&OPS, combined nasopharyngeal and oropharyngeal swab; PWG, purified water gargle; ND, not detected.
Discussion

In this study, we compared the suitability of self-collected PWG versus NPS&OPS collected by healthcare professionals as testing materials for RT-PCR-based IAVs detection in outpatients with febrile respiratory symptoms. The present study demonstrated high overall concordance (85.0%) between paired specimens. Previous studies have reported that testing combined NPS&OPS from one individual has been shown to increase the sensitivity for the detection of respiratory viruses and improve the reliability of the results.\(^2,22\) In our hospital, combined NPS&OPS has been the recommended specimen type for viral detection for many years. This was also the reason why NPS&OPS was chosen as the comparison specimen in this study. However, in contrast to the standardized sampling process of NPS or OPS, the methods of PWG collection were varied, such as gargling 3 times with 3 mL of normal saline,\(^14\) 5 mL of normal saline gargling for 20s,\(^15\) or 30s,\(^17,18\) 10 mL of normal saline for 20s,\(^10\) 20 mL of injection water for 5s,\(^19\) 5 mL of natural spring water for 20s,\(^13\) 5–7 mL of tap water for 30s,\(^12\) 10 mL of tap water for 15s,\(^16\) 1/4 of a glass water for 5 to 20s.\(^5\) Some studies asked patients to swish for few seconds followed by tilting their heads back and gargling for few seconds and repeated swish/gargle cycle.\(^11,13,15\) Previous studies demonstrated that viral RNA in drinking water could be stable at room temperature for at least 3 to 7 days.\(^13,15\) According to these studies, we chose to use 5 mL of commercial purified drinking water, which can effectively sample the entire oropharynx avoiding excessive sample dilution and is easily get without unpleasant saltiness of saline. Patients were required to repeat the swish/gargle cycles when collecting the samples. In this study, PWG was more acceptable than NPS and OPS for patients and could be collected by themselves as required.

In our study, the sensitivity of PWG was lower than those of NPS and OPS (80.1% vs 96.6%). The lower sensitivity of PWG in the current study may be due to the reference sample type selected, assays/kits used, and pathogens detected. Previously, gargle has been reported to show promising performance compared to NPS for the detection of SARS-CoV-2.\(^14,15,23\) A recent meta-analysis including 5922 patients from 16 studies showed that the detection sensitivity of saliva was 83.2% for detecting SARS-CoV-2.\(^24\) In this study, we reported a similar sensitivity in detecting IAVs. However, Dumareq et al showed that the clinical sensitivity of the gargle was 95.3%, similar to the sensitivity (93.8%) of combined NPS&OPS for SARS-CoV-2 detection.\(^13\) Another similar study revealed that the sensitivity of gargle for detecting samples of outpatients was 85.14%.\(^25\) Kohmer et al showed that the sensitivity of gargle was 89.1% among 102 adults with a confirmed SARS-CoV-2 infection within 48 h of initial diagnosis.\(^16\) Goldfarb et al found that self-collected gargle specimens were more sensitive (98% versus 79%) and more acceptable than saliva samples in an outpatient cohort.\(^15\) One small study found gargle to be a more sensitive sample type than throat swab for the detection of respiratory pathogens.\(^20\) One study based on automated molecular point-of-care test revealed that, by combining conventional RT-PCR and sequencing results, the sensitivity for influenza detection in gargle was 97.1%.\(^19\)

In the current study, lower viral loads were observed for the majority of PWG specimens. Compared with 3mL of NPS&OPS, using 5mL of PWG may produce a dilution effect. Theoretically, 3.3 of ΔCt values means a 10-fold dilution, given 100% PCR efficiency.\(^26\) The difference between NPS&OPS and PWG of this study was 4. We believe that the most critical reason is the higher concentration of IAV replicates in the NP/OP regions. Direct sampling from these regions with swabs can collect more viruses. Fifty-nine patients (19.9%) had IAVs detected in their NPS&OPS but not in the PWG. Therefore, for clinically suspected IAV-infected patients with negative PWG results, the NPS&OPS should be tested. Furthermore, at high virus concentrations, viral RNA levels (extrapolated from PCR Ct) were significantly higher in NPS&OPS than in PWG, whereas the difference between NPS&OPS and PWG was not significant at lower virus concentrations. Our results are consistent with the findings of previous studies.\(^13,17,25\) Notably, 10 patients had a higher viral load in their PWG than in the NPS&OPS, who had positive IAV results in PWG but negative results in the NPS&OPS (3.4%).

To improve the sensitivity of gargle, some studies required participants to collect gargle in the morning before eating, drinking or brushing teeth.\(^5,18\) Patients were also required to inhale and cough before gargling to increase the viral loads in the oral cavity.\(^25\) Moreover, using a smaller volume of mouthwash and a longer swish may increase sensitivity. It is worth to find better PWG sampling protocol through comparative research in the future. Another study carried out pre-enrichment strategy for gargle before detection.\(^27\) Furthermore, a more direct solution may be to choose higher sensitivity of molecular assays to circumvent the problem of lower viral load in gargle specimens.\(^1\)

Our study had some limitations. First, only adult patients were included in this study. Further evaluation should be conducted in the pediatric population. Second, only outpatients with an early disease onset were included in this study.
As the viral load may be lower during the later period, further studies should involve hospitalized patients for continuous monitoring. Third, although a standardized order of sampling was chosen to minimize the influence of paired specimens, irritation, such as sneezing when collecting swabs, may still influence the results of self-collected PWG. Fourth, considering that the tissue tropism of other respiratory viruses differs from that of seasonal influenza viruses, the accuracy of PWG for other respiratory viruses should be evaluated in future studies.

In conclusion, we observed that, in this direct head-to-head comparison, although the self-collected PWG specimen offered acceptable performance for IAVs molecular testing, NPS&OPS remained a reliable option. Given the convenience of collection, nonviscous gargles are recommended for viral detection during emergencies or under specific conditions (such as self-testing at home). Furthermore, combining optimization and standardization of gargle specimen collection protocol and using higher sensitivity molecular assays to circumvent the problem of lower viral load will strengthen our recommendation.

Acknowledgments
The authors thank the other clinical and laboratory staff at the First Affiliated Hospital of the Zhejiang University School of Medicine. This research was funded by China National Mega-Projects for Infectious Diseases (grant number 2017ZX10103008).

Disclosure
The authors report no conflicts of interest in this work.

References


