

BMJ Open Accuracy of pH strip testing and pH liquid testing versus standard pH meter of gastric contents in critically ill patients: a diagnostic accuracy study

Kanjanaporn Phukpattanachai,¹ Nutch Praditseree,¹ Smith Skjolaas,¹ Sirilux Klaychaiya,¹ Konlawij Trongtrakul ²

To cite: Phukpattanachai K, Praditseree N, Skjolaas S, *et al.* Accuracy of pH strip testing and pH liquid testing versus standard pH meter of gastric contents in critically ill patients: a diagnostic accuracy study. *BMJ Open* 2024;**14**:e081830. doi:10.1136/bmjopen-2023-081830

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<https://doi.org/10.1136/bmjopen-2023-081830>).

Received 07 November 2023
Accepted 19 June 2024



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

²Internal Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

Correspondence to

Dr Konlawij Trongtrakul;
konlawij@live.com

ABSTRACT

Objective The utilisation of pH level measurements from gastric contents may indicate the preferred tip position of a nasogastric tube or monitor the efficacy of stress ulcer prophylaxis in critically ill patients. We aimed to determine the accuracy of pH strip (pHS) tests and pH liquid (pHL) tests compared with the standard pH meter (pHM).

Design Diagnostic accuracy study.

Setting Gastric contents from medically critically ill patients.

Participants In total, 113 gastric samples were collected from 27 critically ill patients.

Outcome measure The level of pH measured by pHM, pHS and pHL.

Results The pH values measured by pHM, pHS and pHL were 5.83 (IQR 5.12–6.61), 5.50 (IQR 5.00–6.00) and 5.75 (IQR 5.25–6.25), respectively. The pHS test showed greater accuracy, exhibiting a more positive correlation with the standard pHM measurement than the pHL test, with $Y=0.95X+0.56$; $\rho=0.91$, $p<0.001$, and $Y=1.09X-0.72$; $\rho=0.75$, $p<0.001$, respectively. However, the pHS test demonstrated less agreement with the pHM than the pHL test, with biases of -0.27 versus 0.18 , respectively. Noticeably, a slight variation in pHL from the standard pH values was found when we measured gastric contents with a pH lower than 5.

Conclusion Both the pHS and pHL methods were good options for measuring gastric pH in critically ill patients. However, it was advisable to find alternative approaches to the pHL testing method when anticipated gastric acidity levels fall below 5.

Trial registration number TCTR20220530004.

INTRODUCTION

Critically ill patients usually require a nasogastric (NG) tube for feeding and delivering medications. The American Society for Parenteral and Enteral Nutrition stated that each year, more than 245 000 patients required feeding via an NG tube.¹ Although inserting an NG tube is considered a low-risk procedure, misplacement of the tube can lead to severe complications, including oesophageal perforation, aspiration pneumonia,

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ A diagnostic accuracy study was conducted to evaluate the precision of the pH strip test, along with an alternative approach using pH liquid testing, to determine the pH levels of gastric content in critically ill patients.
- ⇒ The pH levels of gastric contents were examined and interpreted by four independent raters using both pH strip and pH liquid test techniques, all of whom were blinded to the pH values obtained from the standard pH meter measurement.
- ⇒ Considering the wide range of gastric pH levels, we collected gastric content samples from both fasting critically ill patients and those receiving stress ulcer prophylaxis.
- ⇒ Given that pH levels represent continuous outcomes, the accuracy of pH strip and pH liquid tests was assessed using regression slope and Bland-Altman plots for bias evaluation.
- ⇒ Stress ulcer prophylaxis may affect the precision of pH testing in determining the position of the nasogastric tube, particularly when gastric pH levels exceed 5, but it was not considered to impair the accuracy of the pH comparisons on testing in our study.

pneumothorax, pulmonary haemorrhage or even death in severe cases.^{2–5}

During the period spanning from September 2011 to March 2016, nearly 100 incidents involving the unintended introduction of fluids and medications into the respiratory tract instead of the gastrointestinal tract resulted in suffocation.^{3 4} Tragically, 35 of these cases were fatal, although the direct correlation attributed to misplaced NG tubes remained inconclusive.^{3 4} For this safety concern, the National Health Service advocates classifying NG tube misplacement as a never-event and entirely preventable with the implementation of systematic protective measures.⁴

It is highly advised to verify the accurate positioning of the NG tube before administering food or medications.⁴ Certain studies suggested using auscultation over the epigastric area to detect the presence of air insufflated into the stomach as a means of confirming the correct placement of the NG tube.^{6,7} Unfortunately, there is still a limitation since gurgling sounds may persist even when the NG tube is incorrectly positioned in the oesophagus or respiratory tract.⁸

Under normal circumstances, the pH levels in an empty stomach are typically around 1–3, increasing to 4–5 after a meal, while endotracheal aspirate usually falls within the range of 6–9.5.^{3,9} Therefore, the application of pH level measurement from the NG tube contents can serve as an indication that the NG tube tip is optimally positioned in the stomach.⁵ Although pH meters (pHMs) are established as the reference standard for assessing gastric pH,⁹ their practicality for use at the bedside is limited. Acknowledging the safety, reliability and feasibility of bedside testing, several studies have recommended employing pH strips (pHS) to ascertain a pH level threshold ranging between 4 and 5.5, serving as indicators of appropriate NG tube placement.^{3,9–11} It is worth noting that a pH below a threshold of 5 has been recognised as the most reliable indicator for confirming NG tube positioning, as indicated by a decision analytical model.¹²

Interestingly, despite one study indicating a favourable correlation between the use of pHS and the pHM method,¹³ the other two studies revealed only a moderate correlation between these two techniques.^{14,15} Moreover, these two studies also highlighted a considerable dispersion between the pHS and pHM measurements.^{14,15} Furthermore, another study revealed constraints in the accurate use of pHS in differentiating between gastric contents and non-gastric contents. Among patients without prior exposure to proton pump inhibitors (PPIs), the accuracy of pHS was limited to 76% and 77% for those with PPI exposure when considering pH levels below 5.5 as the criterion.¹⁰

The current study aims to further enhance our understanding of pH measurement methods. We examined the accuracy of pHS across a wide range of pH results under various gastric content conditions, including both pre-PPI fasting and post-PPI scenarios. Additionally, we explored the utility of pH liquid (pHL) testing as an alternative pH measurement method. These pH assessments may serve to safeguard against the misplacement of NG tubes in critically ill patients.

METHODS

Study design and participants

A diagnostic accuracy study was conducted on gastric samples from critically ill patients who were admitted to the medical intensive care units (ICUs) of the Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand, between June and July 2022. Our team aimed to study the accuracy of pH measurement via pHS testing and pHL

testing in comparison with the standard pHM measurement. Our objective was to evaluate the accuracy of pHS and the alternative testing method, pHL, which might be another option that is practical, simple and available at the bedside. Additionally, we wanted to prove the accuracy of the pHL testing method and whether it is acceptable in clinical practice, as it had not yet been demonstrated.

We adhered to the Standards for Reporting Diagnostic Accuracy Studies (2015) guidelines for study operation and report.

Inclusion and exclusion criteria

The inclusion criteria were critically ill patients who were admitted to the medical ICUs aged 18 or over and received NG tube insertion before study enrolment. Gastric content was retrieved following a minimum 2-hour fasting period. The exclusion criteria were patients who were pregnant, patients with upper gastrointestinal bleeding and those who were unwilling to participate in the study. Our hospital has a policy requiring the use of PPIs as a stress ulcer prophylactic medication for patients who are at high risk for stress-related mucosal damage. Therefore, all of the study participants received this medication.

Gastric content pH measurement

We collected a 5-mL sample of gastric contents from fasting patients as convenient. The gastric contents were tested within 1 hour postcollection by three methods of pH measurement, including pHM, pHS and pHL. The results read from the pHS and pHL tests were interpreted by four independent raters, all of whom tested negative for colour blindness.

The reference gastric pHM was measured by the Eutech pH 700 pH/mV/°C/°F Bench Meter (Eutech Instruments, Singapore), as shown in [figure 1A](#). This machine reports the pH with a scale of two decimal places (100th) ranging from –2 to 16, with an accuracy of ± 0.01 and a resolution of 0.01.¹⁶ The coefficient of variation for pH measurements was less than 0.43%.¹⁷

The pHS testing method was measured by the MQuant pH indicator strips (Merck KGaA, Darmstadt, Germany), as shown in [figure 1B](#). This strip reports pH levels ranging on an integer scale from 0 to 14, interpreted by four different shades of colour per pH level. We estimated it to the nearest 0.5 when the colour was in between two shades of colour.

The pHL testing method was measured by a drop of the V UNIQUE v-color 4590 solution (Better Syndicate, Bangkok, Thailand) and interpreted using colour shading, as shown in [figure 1C](#). The pHL testing reports pH with a scale of one decimal place (10th) ranging from 4.5 to 9. Thus, when the colours were determined to be between two levels of pH, we estimated the results to be in the quartile between decimal places. For example, for colour levels located between 5.5 and 6, we estimated them to be at 5.75.

The pHS testing required gastric contents of slightly less than 1 mL. On the other hand, the pHL testing method

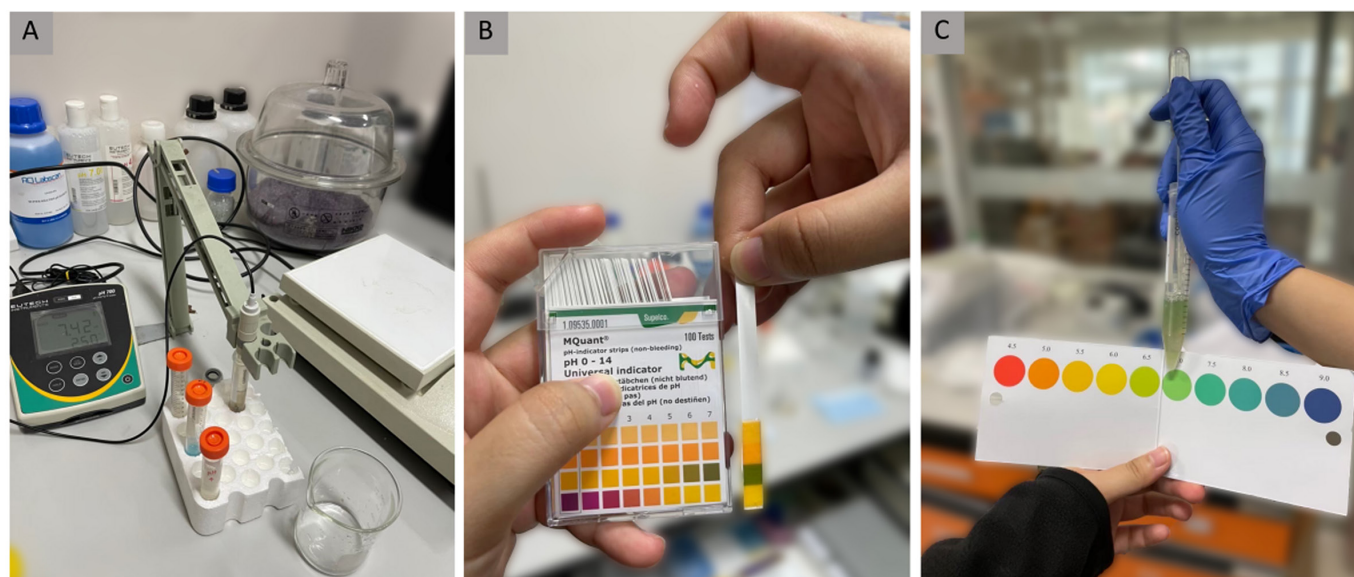


Figure 1 (A) pH meter (reference standard), (B) pH strip testing and (C) pH liquid testing. *Note: The person depicted is not a patient and was taken with the participant's knowledge.

needed at least 5 mL of gastric contents. One researcher took pictures of both testing methods, which were then separately interpreted by independent raters with no known actual pH value from the pHM measurement.

Data collection

We collected data on patients' demographics, including age, sex, body mass index (BMI) and pre-existing comorbidities. ICU admission data were also evaluated, including the presence of sepsis or septic shock, source of infection, organ dysfunction, vital signs, severity of illness measured by the sequential organ assessment (SOFA) score, laboratory investigations and initial management. The dataset could be accessed elsewhere.¹⁸

Sample size

The sample size calculation was based on a given significance level of 0.05, a power of 0.80, a null hypothesis linear regression slope of 0.50 and a projected SD of 1.50 for the pHM method readings. This indicated a need for at least 73 gastric samples to estimate a linear regression slope of 0.90 for the pHS testing method. Similarly, considering a linear regression slope estimate of 0.85 for the pHL testing method, 100 gastric samples were required. Therefore, in total, 112 (111.1) gastric samples were compulsory for the study when considering 10% of the data missing during collection.

Statistical analysis

Continuous data were expressed as median and IQR. Categorical data were reported as numbers and percentages. The reliability of the pHS and pHL tests, to define the extension of the pH testing on its replicability, was determined using the intraclass correlation coefficients (ICC). The ICC reliability value ranges between 0 and 1, with an ICC value approaching 1 indicating excellent reliability.¹⁹ We acquired 50 gastric samples interpreted by four independent raters

to check for the reliability of pHS and pHL measurements based on the single rater/measurement, absolute agreement and two-way random effects model.¹⁹

The value of pH measurement was reported as median and IQR and presented graphically by the violin plot. Comparing three methods for pH measurement, we used the Friedman test and analysed the post hoc pairwise comparison by Dunn's test adjusted with the Bonferroni method.

The linear regression analysis was also used to indicate the linear equations between the two methods of pH testing corresponding to the standard pHM value. The linear equation was displayed as $Y=aX+b$, where Y denotes the predicted value of pHM, a is the slope of the equation, X is the value of pHS or pHL and b is the intercept of the equation. Additionally, Spearman's correlation (ρ) was used to determine the correlation between the pHS versus the pHM and the pHL versus the pHM. The ρ value ranges also between 0 and 1, considering a strong positive correlation when this value is approaching 1.²⁰ Finally, we tested for the agreement of pHS and pHL against the pHM values illustrated by the Bland-Altman plot.

Statistical analyses were performed using the Stata Statistical Software (Release V.17, 2021, College Station, Texas, USA; StataCorp). A p value of <0.05 was considered statistically significant.

Patient and public involvement

Patients and/or the public were not involved in the design, conduct, reporting, or dissemination plans of this research. Refer to the Methods section for further details.

RESULTS

Baseline characteristics

A total of 113 gastric samples were collected from 27 ICU patients. The participants' median age was 71 (IQR

Table 1 Patients' baseline characteristics

Baseline characteristics	Total (n=27)
Age (years)	71 (56–81)
Male, n (%)	13 (48)
Body mass index (kg/m ²)	22.2 (19.2–24.2)
Pre-existing comorbidity, n (%)	
Hypertension	20 (74)
Dyslipidaemia	12 (44)
Diabetes mellitus	8 (30)
Chronic kidney disease	7 (26)
Sepsis/septic shock at ICU admission, n (%)	18 (67)
Source of infection, n (%)	
Respiratory tract	14 (52)
Urinary tract	4 (15)
Gastrointestinal tract	3 (11)
Vital signs at ICU admission	
Body temperature (°C)	37.1 (36.7–38.0)
Mean arterial pressure (mm Hg)	83 (70–98)
Heart rate (beats per minute)	102 (88–127)
Respiratory rate (breaths per minute)	26 (16–32)
Sequential organ failure assessment score	5 (2–9)
Laboratory investigations	
Haemoglobin (g/L)	107 (75–120)
White blood count (10 ⁹ cells/L)	12.3 (9.3–19.0)
Platelet count (10 ⁹ cells/L)	257 (118–336)
Total bilirubin (mg/dL)	0.74 (0.41–0.97)
Creatinine (mg/dL)	1.5 (0.7–2.7)
Serum lactate (mmol/L)	3.7 (2.2–5.6)
Management at ICU admission	
Mechanical ventilation, n (%)	22 (82)
Vasopressor administration, n (%)	7 (26)

Continuous data are presented as median and IQR.
ICU, intensive care unit.

56–81) years, 48% were male and their BMI was 22.2 (IQR 19.2–24.2) kg/m². Hypertension was the most common pre-existing comorbidity (74%). Our cohort contained 67% of patients who were diagnosed with sepsis or septic shock, where the respiratory tract was the main source of infection. The severity of illness measured by SOFA was 5 (IQR 2–9) and the median serum lactate level was 3.7 (IQR 2.2–5.6) mmol/L. Most patients (82%) were supported by mechanical ventilation (table 1).

We found strong reliability in both the pHS and pHL methods, with ICC values of 0.93 (0.90–0.96) and 0.94 (0.90–0.96), respectively, as detailed in online supplemental table 1 and figure 1, with four interpreters assessing pH level readings.

Table 2 Measurements of gastric pH by each technique

Parameters	pH level	P value
pH measurements*		<0.01
pH meter	5.83 (5.12–6.61)	–
pH strip	5.50 (5.00–6.00)	–
pH liquid	5.75 (5.25–6.25)	–
Differences between techniques†		
pH strip and pH meter	–0.27 (–0.49–0.03)	<0.01
pH liquid and pH meter	–0.02 (–0.41–0.43)	0.04
pH liquid and pH strip	0.50 (–0.25–0.75)	<0.01

*Friedman test.

†Pairwise comparison by Dunn's test adjusted with Bonferroni method.

The distribution of gastric pH measurements from three methods is shown in table 2 and online supplemental figure 2. The median values of pHM, pHS and pHL were 5.83 (IQR 5.12–6.61), 5.50 (IQR 5.00–6.00) and 5.75 (IQR 5.25–6.25), respectively, with an overall $p<0.01$ by the Friedman test. There were significant differences between pHS versus pHM ($p<0.01$) and between pHL versus pHM ($p=0.04$), in pairwise comparisons.

The linear function of pHS for estimating pHM was $Y=0.95$ (95% CI 0.88 to 1.01) $X+0.56$ (95% CI 0.18 to 0.93). Additionally, a strong correlation between pHS and pHM was found, with a rho value of 0.91 (95% CI 0.86 to 0.96), $p<0.001$ (figure 2A). When the pHL was provided for determining the pHM, the linear function was $Y=1.09$ (95% CI 0.85 to 1.34) $X-0.72$ (95% CI –2.14 to 0.70), with an optimal rho coefficient at 0.75 (95% CI 0.60 to 0.89), $p<0.001$ (figure 2B).

The Bland-Altman plots depict the differences between the meanings of pHS versus pHM (figure 3A) and pHL versus pHM (figure 3B). The bias values between techniques were –0.27 (95% CI –1.24 to 0.70) and 0.18 (95% CI 1.88 to 2.24), respectively. These plots demonstrated a greater agreement for the pHL testing method with the standard pHM measurement than that of the pHS testing method. Noticeably, a slight variation in pHL from the standard pH values was found when we measured gastric contents with a pH lower than 5. Additionally, online supplemental figure 3A,B display the distribution of results between the pHS test versus the pHM test and the pHL test versus the pHM, respectively.

DISCUSSION

Our study demonstrated that gastric pH testing with the pHS and pHL testing methods was acceptable for substituting the standard pHM measurement. Although the pHS value was slightly lower than that of the pHM (pHS–pHM=–0.27 (IQR –0.49 to 0.03), $p<0.01$), pHS remains capable of determining pHM with a strong correlation ecoefficiency ($Y=0.95$ (95% CI 0.88 to 1.01) $X+0.56$ (95% CI 0.18 to 0.93) and rho=0.91 (95% CI 0.86 to 0.96),

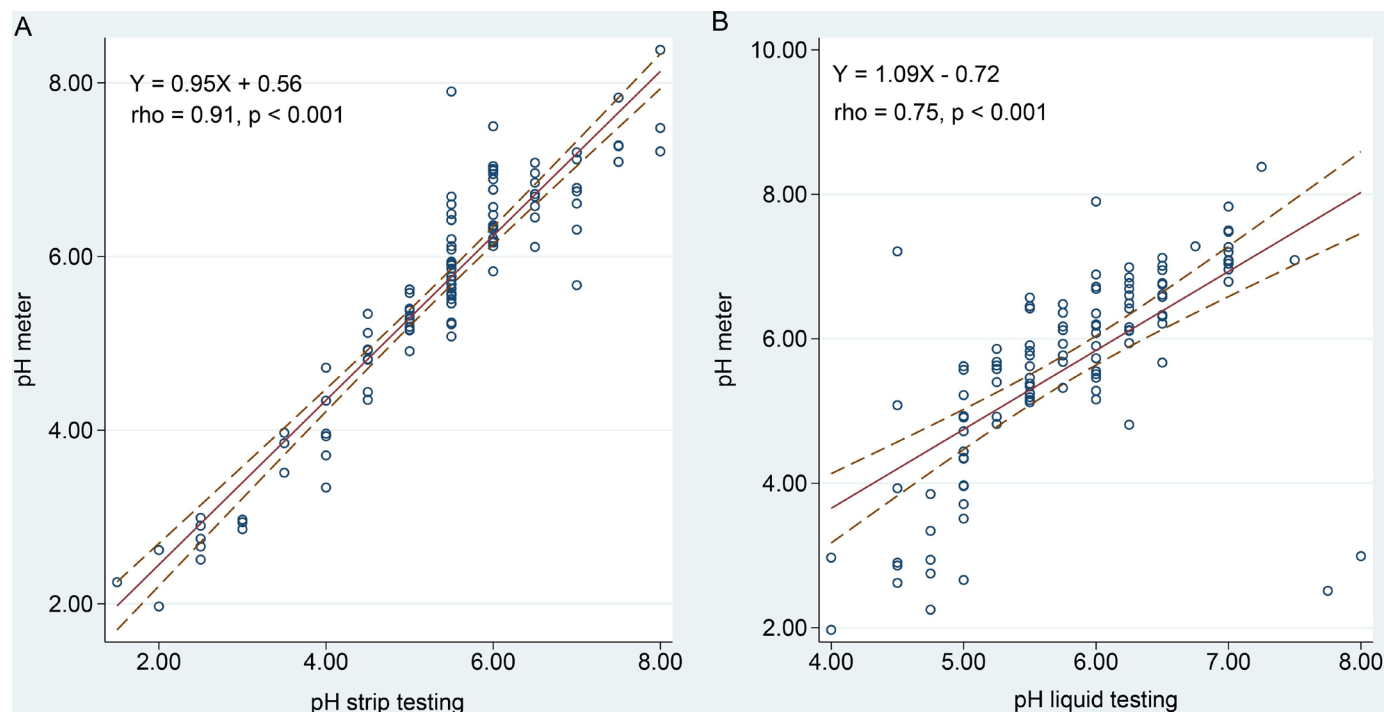


Figure 2 The regression line and Spearman's correlation of (A) pH strip testing versus pH meter and (B) pH liquid testing versus pH meter.

$p < 0.001$, respectively). Fewer differences between the pHL and the standard pHM were found (pHL–pHM=−0.02 (IQR −0.41 to 0.43), $p=0.04$). Unfortunately, the pHL method was less correlated with the pHM method ($Y=1.09$ (95% CI 0.85 to 1.34) $X-0.72$ (95% CI −2.14 to 0.70) and $\rho=0.75$ (95% CI 0.60 to 0.89), $p < 0.001$, respectively). Provided with a strong correlation when the slope of the linear function and the rho coefficient were close to 1, we could then conclude that the pHS testing method was more positively related to the pHM measurement than

that of the pHL testing method for measuring gastric content pH. Unfortunately, the pHS testing method had less agreement with the pHM than that of the pHL testing method (the bias of −0.27 (95% CI −1.24 to 0.70) vs 0.18 (95% CI 1.88 to 2.24), respectively). The reason behind the disagreement might be related to a finer scale of pHL than that of pHS.

There was a paucity of studies that determined the linear function and the correlation between pH testing methods. One study regarding critically ill paediatric

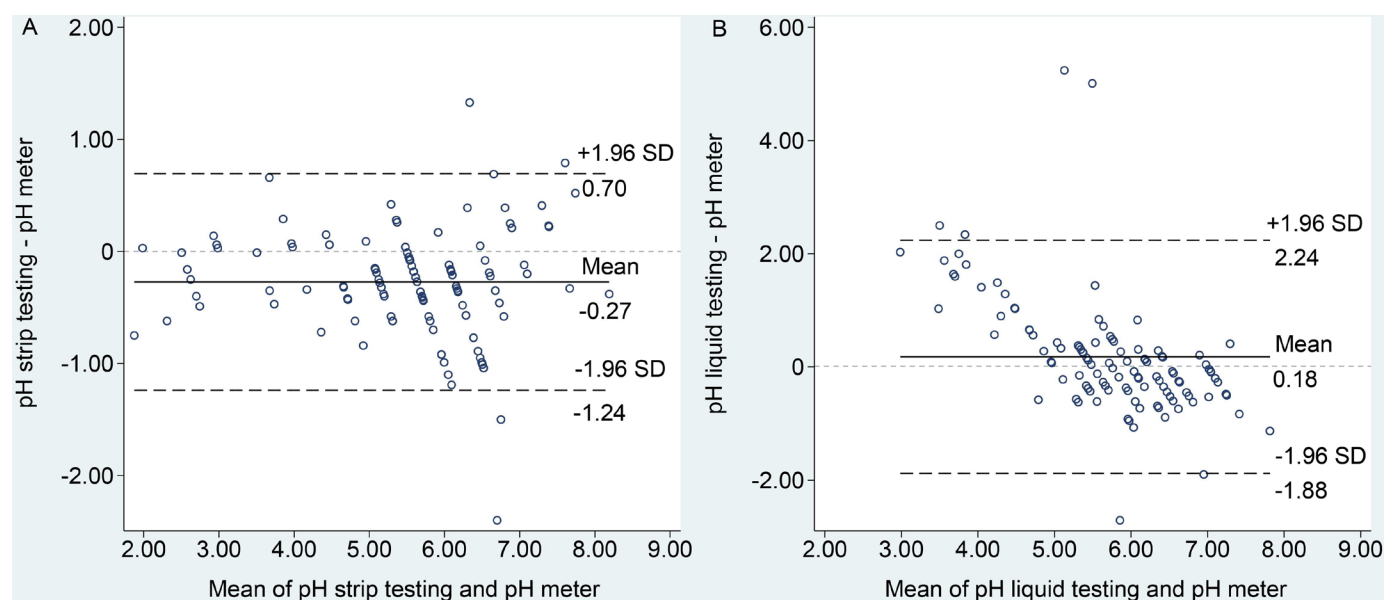


Figure 3 The Bland-Altman plots display the differences versus the mean value of (A) pH strip testing versus pH meter and (B) pH liquid testing versus pH meter.

patients reported a moderate correlation between the pH paper testing and the pHM, with a Pearson correlation of 0.59; $p < 0.001$.¹⁴ The Bland-Altman analysis showed differences between the two methods of 0.41 with a wide dispersion (± 2 SD values equalled -3.68 to 4.88).¹⁴ Another study was reported by Bradley *et al.*¹⁵ They compared gastric pH measurement from pH paper with a hand-held pHM in critically ill surgical patients. A concordance correlation coefficient of 0.896 was demonstrated. The mean difference between these two methods was 0.75, with ± 2 SD values equalling -0.41 to 1.45 reported using the Bland-Altman analysis.¹⁵ Although there were differences between the populations studied, our pHS testing method slightly outperformed these two studies in terms of a less mean bias. Alternatively, in another study, a good agreement between the pH indicator paper and the pHM was found in the anaesthetised patients.¹³ The mean difference between pH measurements was 0.1 and the limits of agreement ranged between -0.1 and 0.3 , according to the Bland-Altman analysis.¹³ A higher performance of pH indicator paper with a finer scale was used in this study.¹³ This might be the reason for the less bias found in this study, where a larger scale of pHS testing was used in ours.

Several studies in the past have demonstrated the accuracy of pH indicator paper in measuring gastric contents. However, using the PHL testing method to analyse gastric contents has not been indicated in prior studies. Although PHL testing delivers less accuracy in terms of linear function and correlation, the PHL test reveals considerably less bias than that of the pHS testing method. One possible reason could be related to the fact that the PHL testing method was interpreted for two decimal places versus one decimal place for the pHS testing method. Interestingly, gastric pH measured by the PHL indicated a slight discrepancy from the standard value when the actual gastric pH was lower than 5.

In our investigation, we encountered two outliers between the PHL and pHM measurements during the Bland-Altman analysis. Initially, these outlier samples were attributed to the presence of bile acid, which led to inaccurately high pH readings when assessed using the PHL method. However, upon closer examination, it was revealed that the pH levels obtained via the PHL method were 7.75 and 8, while the pHM readings were 2.51 and 2.99, respectively. Remarkably, these pHM levels deviated significantly from the typical range of bile acid content pH levels (6.50–8).²¹ Unfortunately, we were unable to account for these discrepancies. One plausible explanation could be linked to the admixture of a high concentration of gastric acid with trace amounts of bile acid, which inherently contains green colour pigments, thus complicating the differentiation of genuinely elevated pH levels (green colour) through the PHL method.

Although both pHS and PHL are optimised for use in clinical practice, we consider pHS testing to be superior. The pHS testing technique promotes more accurate colour perception (using four-colour shadings per pH

level), uses a smaller volume of gastric contents, reports a wider range of pH levels and uses fewer complex procedures at the bedside. Therefore, the pHS test is acceptable for substitution for the pHM and can be used to measure gastric pH in critically ill patient settings to confirm the location of the NG tube.² However, when gastric pH levels exceed 5 or 5.50, especially in cases involving enteral feeding or the administration of medications such as PPIs that elevate gastric pH, further assessment through chest radiography is warranted to ensure the accurate placement of the NG tube.^{2–4} Furthermore, monitoring gastric pH could serve as a valuable tool in monitoring the efficacy of stress ulcer prophylaxis, where maintaining a gastric pH level above 4 is recommended.²²

This research proved that the PHL testing technique was acceptable for measuring gastric pH in critically ill patients. However, some limitations should be considered. For example, the PHL testing method is less reliable when gastric pH levels are less than 5, a larger amount of gastric volume (5 mL) is required, and there is a possibility that the colour of gastric contents, especially bile acid, might interfere with the pH reading.

Relying solely on pHS or PHL tests for measuring gastric pH in the presence of bile acids or blood poses challenges. These substances introduce complexities that can hinder accurate visual interpretation of the pH test due to variations in colour shades. Indeed, it is advisable to exercise caution when using both pHS and PHL when blood is present. In situations where the presence of blood or bile acid is not discernible, additional intricate tests may be required, such as pHM or a combination of pH testing methods offering varied colour interpretations.

There were some limitations in this study. First, the pHS testing and PHL testing methods report a different scale of pH measurement, with an estimated one (10ths) and two (100ths) decimal places, respectively. This might cause a discrepancy between the regression slope or correlation and the Bland-Altman analyses. Conceptually, the regression slope and correlation identify the strength of linear association, while the Bland-Altman analysis is recommended for comparing two methods of quantity measurement.²³ Further studies with a finer scale for pHS testing may be worth investigating. Second, a few of the differences in pH values as measured by pHS and pHM testing methods were outside the limits of agreement. These outliers were mostly related to the colour of gastric contents that interfere with properly reading the correct shade of pH, for instance, the presence of bile acid and blood, among others. One should be aware of the limitations of each pH measurement method and accordingly allocate gastric contents for a more precise method of pH measurement. Finally, stress ulcer prophylaxis may affect the precision of pH testing in determining the position of the NG tube, especially when gastric pH levels exceed 5. However, the use of PPIs was not considered to compromise the accuracy of the pH testing in our study, as it was essential to include gastric samples from a wide range of

pH levels to effectively demonstrate the accuracy of the pH testing methods.

CONCLUSION

Both pHS and pHL testing methods offer the ability to assess gastric pH contents in critically ill patients, providing viable alternatives to pHM measurement. However, our findings revealed greater correlation but less agreement for the pHS testing method compared with the reference pHM measurement, in contrast to the pHL testing method. Furthermore, it was advisable to refrain from using the pHL testing method when anticipated gastric content pH levels are expected to be below 5.

Acknowledgements We sincerely appreciate the ICU nurses of the Maharaj Nakorn Chiang Mai Hospital who obtained gastric content samples. Furthermore, we were grateful to the Department of Biochemistry Faculty of Medicine, Chiang Mai University, who gave us all the necessary equipment to complete the study. Additionally, we would like to acknowledge Mr Jonathan Stream and Ms Laurel Coutts, English Consultants, who assisted in proofreading and editing all essential elements.

Contributors KP, NP, SS and SK prepared the proposal, gathered the patient baseline demographics, collected gastric samples and measured gastric pH, analysed basic statistical information and drafted the manuscript. KT conceptualised the study objective, analysed the advanced statistical analyses, evaluated the results and revised the manuscript for important intellectual content. All authors have read and approved the final version of the manuscript. KT is responsible for the overall content as the guarantor. We used ChatGPT for assistance with proofreading and refining grammatical accuracy.

Funding This study was supported by the Research Fund of the Faculty of Medicine, Chiang Mai University, Thailand, grant number 026-2566. The funder had no roles involved in any steps of the study.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval The ethical approval was obtained from the Research Ethics Committees of the Faculty of Medicine at Chiang Mai University (grant no. MED-2565-08765) on 17 February 2022. Written informed consent was obtained either from the participants or the participant's next of kin. Additionally, this study was registered via <https://www.thaiclinicaltrials.org>, ID no. TCTR20220530004, on 30 May 2022.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. The data presented in this study are available on Mendeley at <https://data.mendeley.com/datasets/rwy35zmygp/1>.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is

properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Konlawij Trongtrakul <http://orcid.org/0000-0002-2544-7322>

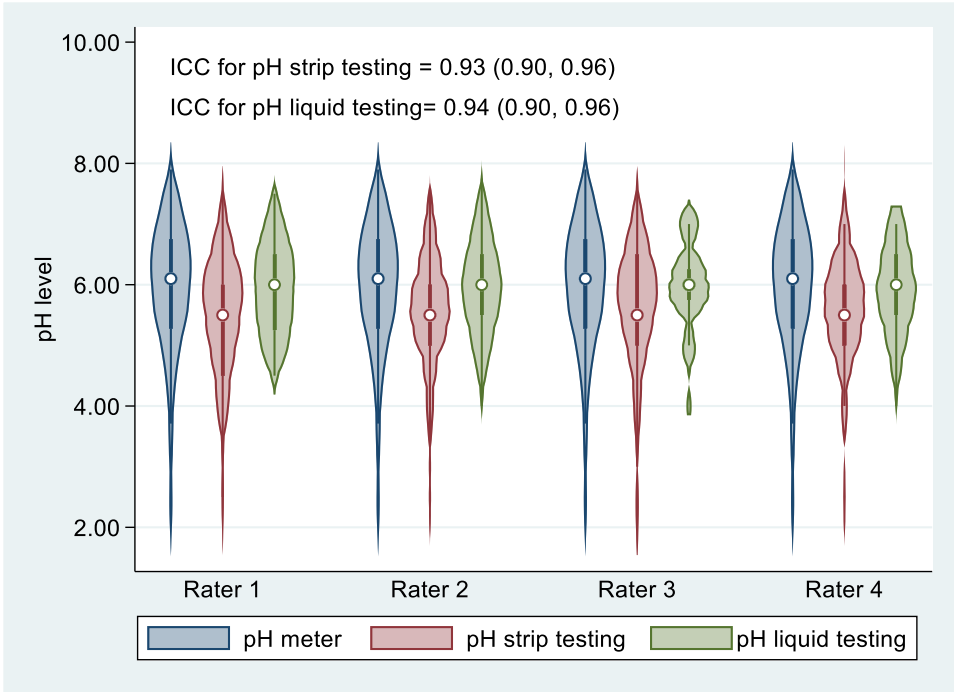
REFERENCES

- Grant MJ, Martin S. Delivery of Enteral nutrition. *AACN Clin Issues* 2000;11:507–16.
- Hardy JP, Ghaye B, Dermesropian F. Hazardous removal of a misplaced Nasogastric tube. *J Belg Soc Radiol* 2020;104:44.
- NHS Improvement. n.d. Resource set initial placement checks for Nasogastric and Orogastric tubes. Available: <https://www.england.nhs.uk/publication/patient-safety-alert-nasogastric-tube-misplacement-continuing-risk-of-death-and-severe-harm>
- NHS Improvement. n.d. Patient safety alert Nhs/PSA/Re/2016/006: Nasogastric tube Misplacement: continuing risk of death and severe harm. Available: <https://www.england.nhs.uk/publication/patient-safety-alert-nasogastric-tube-misplacement-continuing-risk-of-death-and-severe-harm>
- Metheny NA, Krieger MM, Healey F, et al. A review of guidelines to distinguish between gastric and pulmonary placement of Nasogastric tubes. *Heart Lung* 2019;48:226–35.
- Initial and ongoing verification of feeding tube placement in adults (applies to blind insertions and placements with an electromagnetic device). *Crit Care Nurse* 2016;2:e8–13.
- Bourgault AM, Halm MA. Feeding tube placement in adults: safe verification method for blindly inserted tubes. *Am J Crit Care* 2009;18:73–6.
- Shlamovits G, Shah N. Nasogastric Intubation, Available: <https://emedicine.medscape.com/article/80925-overview>
- Gilbertson HR, Rogers EJ, Ukoumunne OC. Determination of a practical pH cutoff level for reliable confirmation of Nasogastric tube placement. *JPEN J Parenter Enteral Nutr* 2011;35:540–4.
- Rowat AM, Graham C, Dennis M. Study to determine the likely accuracy of pH testing to confirm Nasogastric tube placement. *BMJ Open Gastroenterol* 2018;5:e000211.
- Ni MZ, Priest OH, Phillips LD, et al. Risks of using bedside tests to verify Nasogastric tube position in adult patients. *EMJ Gastroenterol* 2014;49–56.
- Ni MZ, Huddy JR, Priest OH, et al. Selecting pH cut-offs for the safe verification of Nasogastric feeding tube placement: A decision Analytical Modelling approach. *BMJ Open* 2017;7:e018128.
- Søreide E, Søreide JA, Holst-Larsen H, et al. Studies of gastric content: comparison of two methods. *Br J Anaesth* 1993;70:360–2.
- Solana MJ, López-Herce J, Sánchez C, et al. Comparison of continuous pH-meter and intermittent pH paper monitoring of gastric pH in critically ill children. *Eur J Gastroenterol Hepatol* 2012;24:33–6.
- Bradley JS, Phillips JO, Cavanaugh JE, et al. Clinical utility of pH paper versus pH meter in the measurement of critical gastric pH in stress ulcer prophylaxis. *Crit Care Med* 1998;26:1905–9.
- Eutech Ph 700 Benchtop Ph Meter, Available: http://www.eutechinst.com/brochures/bench_meters/eutech_bench_700_series_family_brochure_r1.pdf
- Moroff G, Seetharaman S, Kurtz J, et al. Characterizing the variation in pH measurements with Apheresis platelets. *Transfusion* 2011;51:2377–83.
- Phukpattanachai K, Praditseree N, Skjolass S, et al. Data from: The Accuracy of Ph Strip Testing and Ph Liquid Testing Versus Standard Ph Meter of Gastric Contents in Critically Ill Patients. Mendeley Data, 2023. Available: <https://data.mendeley.com/datasets/rwy35zmygp/1>
- Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155–63.
- Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012;24:69–71.
- Hofmann AF, Mysels KJ. Bile acid solubility and precipitation in vitro and in vivo: the role of conjugation, pH, and Ca²⁺ ions. *J Lipid Res* 1992;33:617–26.
- Madsen KR, Lorentzen K, Clausen N, et al. Guideline for stress ulcer prophylaxis in the intensive care unit. *Dan Med J* 2014;61:C4811.
- Mansournia MA, Waters R, Nazemipour M, et al. Bland-Altman methods for comparing methods of measurement and response to criticisms. *Glob Epidemiol* 2021;3:100045.

The Accuracy of pH Strip Testing and pH Liquid Testing Versus Standard pH Meter of Gastric Contents in Critically Ill Patients: A Diagnostic Accuracy Study

Supplementary Information

Supplemental Fig. 1 and Supplemental Table 1 summarize the measurement of gastric pH levels according to the independent raters’ interpretations. For the pH strip testing method, there were no differences in pH level from each rater ($p=0.55$). Rater number 1 reported pH levels of 5.50 (IQR 4.50, 6.00), while raters’ number 2 to number 4 reported the same pH levels of 5.50 (IQR 5.00, 6.00). This figure was also found in the pH liquid testing method ($p=0.38$). Rater number 1 and number 4 reported pH levels of 5.75 (IQR 5.25, 6.50), while rater number 2 and number 3 reported pH levels of 5.75 (IQR 5.25,6.25). The reliability tested by the intraclass correlation coefficients (ICC) of the pH strip testing and the pH liquid test method was 0.93 (95%CI, 0.90, 0.96) and 0.94 (95%CI, 0.90, 0.96), respectively, which indicates excellent performance for both techniques.



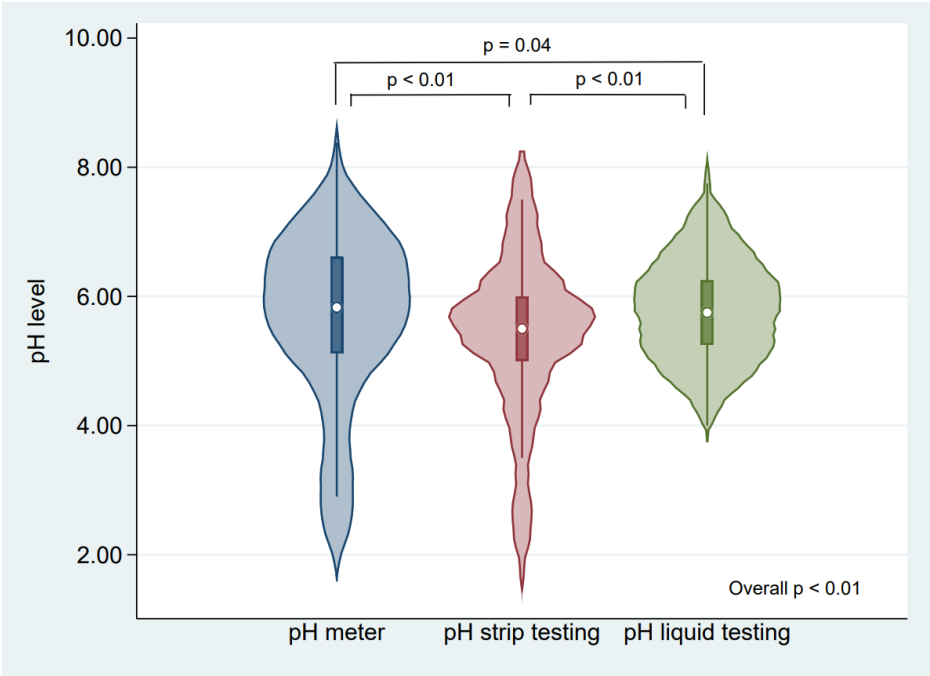
Supplementary Fig. 1 Levels of gastric pH by each method according to independent raters interpretations.

Supplementary Table 1 Measurement of gastric pH according to the independent raters’ interpretations.

Measurement	Rater	pH level	p-value*
pH meter	-	5.83 (5.12, 6.61)	-
pH strip test	Rater 1	5.50 (4.50, 6.00)	0.55
	Rater 2	5.50 (5.00, 6.00)	
	Rater 3	5.50 (5.00, 6.00)	
	Rater 4	5.50 (5.00, 6.00)	
pH liquid test	Rater 1	5.75 (5.25, 6.50)	0.38
	Rater 2	5.75 (5.25,6.25)	
	Rater 3	5.75 (5.25,6.25)	
	Rater 4	5.75 (5.25, 6.50)	

*The p-value was calculated from Friedman statistics.

Supplemental Fig. 2 Distribution of gastric pH measurement comparing three methods.



Supplementary Fig. 3 Histogram of the distribution of gastric pH level according to the pHM test and the pHL test vs. the pHM test.

