PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Factors influencing the diagnostic accuracy of the rapid influenza antigen detection test (RIADT) : a cross-sectional study.
AUTHORS	Tanei, Mika ; Yokokawa, Hirohide; Murai, Kenji; Sakamoto, Rino; Amari, Yu; Boku, Soushin; Inui, Akihiro; Fujibayashi, Kazutoshi; Uehara, Yuki; Isonuma, Hiroshi; Kikuchi, Ken; Naito, Toshio

VERSION 1 - REVIEW

REVIEWER	Gerald H. Stein, MD, FACP Clinical Assistant Professor Department of Medicine University of Florida Gainesville Florida United States
REVIEW RETURNED	02-Oct-2013

GENERAL COMMENTS	Page 1 Title and Key Words; Suggest delete 'immunochromatography,' not of interest to the general practitioner. Page 2: Contribution statement: delete ',3)' Page 3 Typing error: 'Abstract' Suggest authors conform to usual BMJ Abstract sections-see current published articles. P3 L17 & 20 Best to use US or UK spelling throughout: 'center.' Prefer to use VRV throughout manuscript, delete RV+ since the + may be confused with '+' for positive. Prefer delete GS+ & GS-, use VRV+ and VRV-, throughout manuscript to simplify abbreviations. Page 5 L 27, suggest 'supposed to be'(casual English) and use 'may be a helpful.'(formal English) Page 5 L 41 Please spell out RT-CR first time used. Page 7, Please provided the patient selection process, such as all consecutive adult patients with flu-like symptoms or some random process. Page 7, Methods, and Table 2, suggest delete 'pharyngeal erythema,- a subjective examination sign, not symptom. Page 5 L 41 and P7 L 32 suggest delete RV+ system and 'RV+ test.' Page 8 L 32and L36, suggest use 'Verigene test' throughout. P 8 L26 suggest adding brief summary of the immunochromatography method and reporting. Page 8-9, How are the tests results reported? Who performed the tests and reported the results, laboratory technicians, nurses or attending physicians? What was the duration for the VRV report? For Table 1& 3, suggest use 'sex' and delete 'gender.' For Table 2 L 7, suggest 'Verigene test-VRV.' For Table 4, suggest delete 'body' and combine into 3 temperature groups(37.5-38.0).

	P 16 L3-10 might be placed in Methods(P8 L 28). The viral shedding aspect might be expanded in relation to the Testa method causing low sensitivity. P17 L 3 Please support patients were healthy with a statement in Methods section and that patients with co-morbid diseases were excluded if this was done. P17 L 8, consider adding a statement that pediatric group where excluded, the group with the highest sensitivity to RIADT P17 L 29. Please clarify the VRV low specificity for influenza A in relation to the RIADT. Perhaps adding statements in the Abstract conclusion, Key messages and Conclusion that the high specificity of the RIADT of the positive result gives the practitioner firm support for the diagnosis of influenza. Practitioners might be interested that the flu vaccination failed in many cases to prevent flu. References. Suggest following manuscript formatting [http://group.bmj.com/products/journals/instructions-for-authors/formatting/]

REVIEWER	**Carlos Grijalva** **Vanderbilt University, USA**
REVIEW RETURNED	16-Oct-2013

GENERAL COMMENTS	The article by Tanei et al describes factors influencing the diagnostic accuracy of a rapid influenza antigen detection test (RIADT). This was a single center cross-sectional study of 82 outpatient adults presenting with flu-like symptoms, and conducted in Japan during the 2010-2011 season. The investigators used a fully automated virus nuclear acid test (i.e. Verigene) as the reference for estimation of RIADT diagnostic characteristics. Relative to Verigene determinations, the estimated sensitivity, specificity, positive and negative predictive value of RIADT were 72.9%, 91.3%, 95.6% and 56.8%, respectively. The study noted that short time from symptom onset to testing was associated with false positive RIADT results and suggested that chills could be a marker for influenza infection if the RIADT result was negative. I think this is an interesting report on a topic that has been widely discussed. The diagnostic performance of rapid influenza antigen detection tests has been evaluated in multiple articles and reviews, with a largely consistent observation that the sensitivity of the tests seems to be suboptimal although most rapid tests had high specificity relative to culture and/or PCR. I think the current study requires several clarifications. It is necessary to characterize the study selection criteria in detail, so that the reader will be able to appreciate the characteristics of the study population. The study needs to clarify what were the selection criteria for testing, was the testing done systematically, at random or at the discretion of the attending provider? It is unclear why there were only 82 patients in the sample? Were tests done in other subjects not included in the 82 described in the report, if so how they compare with those included in the report? How were these patients selected? Was this number considered enough to support the diagnostic evaluation of RIADT (i.e. provided enough statistical power)? Please provide additional details about the collection of symptoms and body temperature measurements – how often

	recorded and when? Was use of antipyretics considered? Specific comments/suggestions: Abstract – suggest clarify whether testing was systematic or random or what criteria were used for selection of subjects. Please clarify the timing and frequency of symptoms measurements. Please provide 95% confidence intervals for all estimations, so the reader could appreciate the precision of the determinations Key message – suggest clarify what is the clinical implication of the suboptimal sensitivity of the test (e.g. additional tests/evaluations may be required if negative during influenza season) Methods – please clarify the selection criteria for the study and the process to select patient for testing, please clarify if patients fulfilling selection criteria were systematically tested or whether other selection process was used Please clarify the timing and frequency of symptoms measurements and vaccination history Please clarify the activity of influenza during the time when influenza testing was evaluated. It would be useful to describe whether testing was done when there was high or low activity, as this has implications for the interpretation of test results. Please describe how personnel were trained for sample collection and for reading of results from the Verigene system. Some times tests that rely on visual interpretation of findings can provide discordant readings and assessments of reliability are useful. Please clarify if the integrity of the collected samples was evaluated – several studies use a house-keeping gen (e.g. RNP) to verify this Please clarify if the physician attending the patient decided on the use of the Verigene referent (page 8)?– this is very important as selection of certain patients may affect the estimation of tests characteristics (e.g. selection of sicker patients for testing) It is unclear why there were only 82 patients in the sample? Were tests done in subjects not included in the 82 described in the report, if so how they compare with those included in the report? How were these patients selected? Was this number considered enough to support the diagnostic evaluation of RIADT (i.e. provided enough statistical power)? Results Please provide 95% confidence intervals for all estimations, so the reader could appreciate the precision of the determinations In tables 3 & 4, because of the small sample size and apparent distributions, suggest use non-parametric tests for comparison of continuous variables. Please clarify timing of the measurements Discussion Please provide references to support the statement that PCR for influenza is rarely used clinically in Japan – since antivirals appear to be used frequently, it would be useful to clarify if rapid antigen

	detection is widely used to decide therapy in Japan Please provide a reference for the values presented in the second paragraph about the performance of RIADT and clarify what was the test used as reference. Please clarify the correlation of chills with fever – which could be more objective. Clarifying the timing and frequency of the measurements (e.g. at home when symptoms started, or at the clinic after antipyretics were taken) is very important for a proper interpretation of these findings. When describing sensitivity, specificity, etc it is useful to describe the test used as reference. Please provide this information for the descriptions of Verigene diagnostic characteristics I think the conclusions should address the issue of limited sensitivity first, as this could have direct implications for interpretation of testing and to decide whether additional tests may be needed. I would suggest using the STARD statement for reporting of diagnostic tests characteristics (rather than STROBE)

VERSION 1 – AUTHOR RESPONSE

Reviewer Name Gerald H. Stein, MD, FACP

Page 1 Title and Key Words; Suggest delete 'immunochromatography,' not of interest to the general practitioner.

Response: We deleted ‘immunochromatography’ from the title and key word list, as follows:

Factors influencing the diagnostic accuracy of the rapid influenza antigen detection test (RIADT) : a cross-sectional study.

Page 2: Contribution statement: delete ',3)'

Response: We deleted ‘3)’ from contribution statement.

Toshio Naito: 1), 2)

Page 3 Typing error: 'Abstract' Suggest authors conform to usual BMJ Abstract sections-see current published articles.

Response: Thank you for your suggestion. We revised the manuscript to conform to the usual BMJ Abstract format, and corrected the spelling error.

ABSTRACT

P3 L17 & 20 Best to use US or UK spelling throughout: 'center.'

Response: We revised to use US standard spelling throughout, as far as possible (‘center’).

[Design] Single-center, cross-sectional study.
[Setting] Primary-care center, Tokyo, Japan.

Prefer to use VRV throughout manuscript, delete RV+ since the + may be confused with '+' for positive.

Response: We deleted RV+ and replaced it with VRV throughout.

The RIADT and a fully automated respiratory virus nucleic acid test (Verigene Respiratory Virus Plus; VRV), the latter being the gold standard, were performed.

Prefer delete GS+ & GS-, use VRV+ and VRV-, throughout manuscript to simplify abbreviations.

Response: We have revised the manuscript accordingly.

Patients were divided into 4 groups: False Negative (FN), RIADT- and VRV+; True Positive (TP), RIADT+ and VRV+; True Negative (TN), RIADT- and VRV-; and False Positive (FP), RIADT+ and VRV-.

Page 5 L 27, suggest 'supposed to be'(casual English) and use 'may be a helpful.'(formal English)

Response: We have revised accordingly.

The presence of high fever and chills may be helpful indicators of influenza, even if the RIADT result is negative.

Page 5 L 41 Please spell out RT-PCR first time used.

Response: We added ‘reverse transcriptase-polymerase chain reaction’ before ‘RT-PCR.’

Second, the reference or “gold” standard in this study was not viral culture or reverse transcriptase-polymerase chain reaction (RT-PCR), but the Verigene System VRV, which detects influenza virus nucleic acid.

Page 7, Please provided the patient selection process, such as all consecutive adult patients with flu-like symptoms or some random process.

Response: We provided the selection process in the revised METHODS section.

From December 2010 to April 2011, during the influenza epidemic season in Japan,[21] participants were enrolled in the Departments of General Medicine of Juntendo University Hospital and Juntendo University Nerima Hospital, both in Tokyo, Japan. Enrolled were consecutive cases who met the following inclusion criteria: adult patients presenting with any upper respiratory symptoms, and fever ≧37 °C at any time after symptom onset.

Page 7, Methods, and Table 2, suggest delete 'pharyngeal erythema,- a subjective examination sign, not symptom.

Response: We believe your suggestion is correct. We deleted ‘pharyngeal erythema.’

symptoms (sore throat, arthralgia and/or myalgia, headache, chills, cough and/or throat phlegm, and nasal discharge)

Page 5 L 41 and P7 L 32 suggest delete RV+ system and 'RV+ test.'

Response: We deleted RV+ and RV+ test and have used VRV instead.

Page 8 L 32and L36, suggest use 'Verigene test' throughout.

Response: We revised the text to use ‘Verigene test’ throughout.

Specimens were tested by Verigene test according to the manufacturer’s instructions by the physician attending the study patients.
The extraction tray, amplification tray and test cartridge were then loaded onto the Verigene System.

P 8 L26 suggest adding brief summary of the immunochromatography method and reporting.

Response: We added a description of the immunochromatography method, as follows:

RIADTs are immunoassays using the antigen-antibody reaction, based on colloidal gold immunochromatography. The test results are checked visually. The RIADT used for this study, the RapidTesta FLU II, requires a 1×105 tissue-culture infective dose (TCID)50 / mL for type A influenza, and 1.2×105 (TCID)50/mL for type B influenza, to produce a positive result.[22]

Page 8-9, How are the tests results reported? Who performed the tests and reported the results, laboratory technicians, nurses or attending physicians? What was the duration for the VRV report?

Response: The result of VRV was displayed on the reader’s monitor. RIADT was performed by the outpatient doctors or residents, who were experienced with the required skills, and VRV was performed by 5 physicians, who were well trained in the procedure before this research started. We added the following statements to the METHODS section:

(RIADT) When influenza A or B is present, an additional red line appears next to the control red line on the letter ‘A’ or ‘B’ indicated on the test device. The procedures were performed by outpatient physicians and residents who had been well-trained in the technique.
(VRV) The result for each virus type, ‘Detected’ or ‘Not detected,’ was displayed on a monitor. Approximately 2.5 hours was required from sample procurement to final readout.
Specimen collection was performed by one of 5 physicians, who also read the results. These physicians were trained by an instructor from the manufacturer before the study commenced.

For Table 1& 3, suggest use 'sex' and delete 'gender.'

Response: We revised the text to use ‘sex’ instead of ‘gender’.

Male sex

For Table 2 L 7, suggest 'Verigene test-VRV.'

Response: We corrected this to ‘Verigene test-VRV.’

Verigene test-VRV

For Table 4, suggest delete 'body' and combine into 3 temperature groups (37.5-38.0).

Response: We have revised the manuscript as suggested:

Temperature (°C)
Mean
Temperature (°C)
T≦37.4°
T=37.5 to 38.0
T≧38.1

P 16 L3-10 might be placed in Methods (P8 L 28). The viral shedding aspect might be expanded in relation to the Testa method causing low sensitivity.

Response: We understand your suggestion. The statement about the RIADT method was moved to the ‘METHODS’ section. We would like to discuss the reason why RIADT has false negative results early in the ‘DISCUSSION’ section, so the statement about viral shedding was not moved.

METHODS

1. RIADT

RIADTs are immunoassays using the antigen-antibody reaction, based on colloidal gold immunochromatography. The test results are checked visually. The RIADT used for this study, the RapidTesta FLU II, requires a 1×105 tissue-culture infective dose (TCID)50 per mL for type A influenza, and 1.2×105 TCID50/mL for type B influenza, to produce a positive result.[22]

DISCUSSION

The RIADT used for this study, the RapidTesta FLU II, requires a 1×105 tissue-culture infective dose (TCID)50/mL for type A influenza, and 1.2×105 TCID50/mL for type B influenza, to produce a positive result. The influenza virus proliferates in respiratory tract epithelial cells, and appears in respiratory secretions 24 hours before symptom onset. The peak of viral shedding is 24 hours after symptom onset, and then the virus load decreases rapidly.[28] In the current study, we assumed that the amount of virus in the FN group was less than that in the TP group.

P17 L 3 Please support patients were healthy with a statement in Methods section and that patients with co-morbid diseases were excluded if this was done.

Response: We did not exclude older patients or patients with co-morbid diseases. The 82

subjects enrolled in the current study were relatively young and healthy. This is because our hospitals are located in central Tokyo and most patients are young or middle-aged healthy workers. We did not exclude elderly patients or patients with co-morbid diseases. The participants were relatively young and healthy. We have revised the manuscript as follows:

The 82 subjects enrolled in the current study were relatively young and healthy. This is because our hospitals are located in central Tokyo and most patients are young or middle-aged healthy workers. We did not exclude elderly patients or patients with co-morbid diseases. Elderly people tend to present with atypical symptoms and often have underlying primary illnesses.

P17 L 8, consider adding a statement that pediatric group where excluded, the group with the highest sensitivity to RIADT

Response: We added a statement that the participants were all adults and the pediatric group was excluded, as follows:

Also, because almost all of our patients are adults, pediatric patients were not included in this study. This group has distinctive symptoms, and the RIADT has the highest sensitivity in these patients. For this reason, applying similar research methods to different age groups may produce different results.

P17 L 29. Please clarify the VRV low specificity for influenza A in relation to the RIADT.

We found a new reference of VRV [32]. According to the reference, the specificity was 93.2% (95% CI, 91.9%-94.8%). The gold standards are direct fluorescent antibody identification and viral culture. There were a total of 48 samples for which the gold standard was negative but the Verigene test was positive. But 44 of these were positive by sequencing and 4 were negative by sequencing. 93.2% is a relatively low specificity, but if sequencing only was used as the gold standard, the resulting specificity would be higher than 93.2%. We revised the manuscript, as follows:

When direct fluorescent antibody identification and viral culture were used as the gold standard, the sensitivity of this system for influenza A is 98.7% (95% CI, 96.8-99.5%) and the specificity is 93.2% (95% CI, 91.1-99.9%). For influenza B, the sensitivity is 100% (95% CI, 91.8-100%) and the specificity is 99.7% (95% CI, 99.1-99.9%). [32]

Perhaps adding statements in the Abstract conclusion, Key messages and Conclusion that the high specificity of the RIADT of the positive result gives the practitioner firm support for the diagnosis of influenza.

Response: As you suggest, we confirmed the high specificity of RIADT. We added statements about the RIADT's high specificity, as follows:

ABSTRACT [Conclusions] The RIADT sensitivity was low, due to early administration of the test. In the epidemic season, the RIADT should not be used for suspected influenza until 12 hours after symptom onset. A positive RIADT firmly supports the influenza diagnosis; a negative result does not confirm its absence. High fever and chills might indicate influenza, but additional tests are sometimes necessary.

Key messages

The high specificity of the RIADT means that a positive result provides firm support for the diagnosis

of influenza.

CONCLUSIONS

Consistent with previous reports, the sensitivity of the RIADT used in this study was low, due to early administration of the test. Administration of an RIADT too early after symptom onset causes false negative results. In the influenza epidemic season, practitioners should not use RIADT for patients with upper respiratory symptoms and high fever for at least 12 hours after onset. A positive RIADT result after this gives the physician firm support for a diagnosis of influenza. A negative RIADT result does not mean 'no influenza'. Presence of high fever and chills might predict influenza, but additional tests are necessary for patients with specific symptoms inconsistent with a diagnosis of influenza virus infection.

Practitioners might be interested that the flu vaccination failed in many cases to prevent flu.

Response: As you suggest, about 46.6% of influenza patients in our study (27/58, FN group + TP group) had received an influenza vaccination in that season. We are also interested in this limitation of vaccination.

References. Suggest following manuscript formatting [http://group.bmj.com/products/journals/instructions-for-authors/formatting/]

Response: We revised to follow the correct formatting for references.

I enjoyed reviewing this timely important article as a general internist. Its publication before the flu season would be beneficial to many international practitioners.

I am curious if the VRV test reported any Respiratory Syncytial Viruses.

When direct fluorescent antibody identification and viral culture were used as the gold standards, the sensitivity of the VRV system for Respiratory Syncytial Virus was 97.2% (95% CI, 92.1-99.0%) and the specificity was 99.5% (95% CI, 98.7-99.8%).[32]

Reviewer Name Carlos Grijalva
Institution and Country Vanderbilt University, USA

The article by Tanei et al describes factors influencing the diagnostic accuracy of a rapid influenza antigen detection test (RIADT). This was a single center cross-sectional study of 82 outpatient adults presenting with flu-like symptoms, and conducted in Japan during the 2010-2011 season. The investigators used a fully automated virus nuclear acid test (i.e. Verigene) as the reference for estimation of RIADT diagnostic characteristics. Relative to Verigene determinations, the estimated sensitivity, specificity, positive and negative predictive value of RIADT were 72.9%, 91.3%, 95.6% and 56.8%, respectively. The study noted that short time from symptom onset to testing was associated with false positive RIADT results and suggested that chills could be a marker for influenza infection if

the RIADT result was negative. I think this is an interesting report on a topic that has been widely discussed. The diagnostic performance of rapid influenza antigen detection tests has been evaluated in multiple articles and reviews, with a largely consistent observation that the sensitivity of the tests seems to be suboptimal although most rapid tests had high specificity relative to culture and/or PCR. I think the current study requires several clarifications.

It is necessary to characterize the study selection criteria in detail, so that the reader will be able to appreciate the characteristics of the study population. The study needs to clarify what were the selection criteria for testing, was the testing done systematically, at random or at the discretion of the attending provider?

Response: We added the inclusion criteria: adult patients who have any upper respiratory symptoms and fever ≧37℃ at any time after symptom onset. Participants were all consecutive patients. There was no discretion on the part of the attending provider.

From December 2010 to April 2011, during the influenza epidemic season in Japan,[21] participants were enrolled in the Departments of General Medicine of Juntendo University Hospital and Juntendo University Nerima Hospital, both in Tokyo, Japan. Enrolled were consecutive cases who met the following inclusion criteria: adult patients presenting with any upper respiratory symptoms, and fever ≧37 °C at any time after symptom onset.

It is unclear why there were only 82 patients in the sample? Were tests done in other subjects not included in the 82 described in the report, if so how they compare with those included in the report? How were these patients selected?

Response: We enrolled only 82 patients in this study. There were no other subjects who were tested by the RIADT and the Verigene test. We revised the manuscript as follows:

RESULT
A total of 82 consecutive patients meeting eligibility criteria were enrolled from December 2010 to April 2011 (Juntendo University Hospital: 37 patients; Juntendo University Nerima Hospital: 45 patients). There was no selection discretion on the part of the attending physicians.

Was this number considered enough to support the diagnostic evaluation of RIADT (i.e. provided enough statistical power)?

Response: If we detected a significant difference about symptoms between FN group and TN group, for α=0.05, power=0.80, the calculated required sample size is: FN group=67 and TN group=67. We enrolled a total of 82 patients in this study. We plan to conduct a further study of this topic, with more enrolled patients.

Please provide additional details about the collection of symptoms and body temperature measurements – how often recorded and when? Was use of antipyretics considered?

Response: The information about symptoms was collected when an outpatient physician examined the participant. The physician asked him/her if he/she had any symptoms ‘now’, and took his/her temperature. Then the physician filled out the recording document. Only a few participants had used antipyretics and their temperature was around 36°C, but we could not analyze the data with regard to antipyretic use, because we asked about symptoms ‘now.’

METHODS

All of the participants were physically examined and historical data, including age, sex, vaccination status, temperature and symptoms (sore throat, arthralgia and/or myalgia, headache, chills, cough and/or throat phlegm, and nasal discharge) were recorded, as was the time to test from symptom onset. Vaccination status indicated whether an influenza vaccine had been administered during that season before symptom onset. The temperature was taken on presentation by an outpatient physician. Symptoms recorded were those participants reported on presentation. Only a few patients had taken antipyretics and their temperature was around 36°C, but we could not analyze the data with regard to antipyretics use.

Specific comments/suggestions:

Abstract – suggest clarify whether testing was systematic or random or what criteria were used for selection of subjects.

Response: We added a statement about the participants, as follows:

82 consecutive outpatients presenting with upper respiratory symptoms and fever ≧37°C at any time from symptom onset, between December 2010 and April 2011.

Please clarify the timing and frequency of symptoms measurements.

Response: We added a statement about the timing of symptom measurements.

METHOD

Symptoms recorded were those participants reported on presentation.

Please provide 95% confidence intervals for all estimations, so the reader could appreciate the precision of the determinations

Response: We added 95% confidence intervals to the revised manuscript, as follows:

RIADT sensitivity, specificity, positive predictive value and negative predictive value were 72.9% (95% CI, 61.5-84.2), 91.3% (79.7-102.8), 95.6% (89.5-101.6) and 56.8% (40.8-72.7), respectively.

Key message – suggest clarify what is the clinical implication of the suboptimal sensitivity of the test (e.g. additional tests/evaluations may be required if negative during influenza season)

Response: We added a statement about the clinical implication of the low sensitivity of RIADT.

λ The presence of high fever and chills may be helpful indicators of influenza, even if the RIADT result is negative. But additional examination is necessary for patients with symptoms inconsistent with influenza virus infection.

Methods – please clarify the selection criteria for the study and the process to select patient for testing, please clarify if patients fulfilling selection criteria were systematically tested or whether other

selection process was used

Response: We added inclusion criteria and how the participants were recruited, as follows:

From December 2010 to April 2011, during the influenza epidemic season in Japan,[21] participants were enrolled in the Departments of General Medicine of Juntendo University Hospital and Juntendo University Nerima Hospital, both in Tokyo, Japan. Enrolled were consecutive cases who met the following inclusion criteria: adult patients presenting with any upper respiratory symptoms, and fever ≧37 °C at any time after symptom onset.

Please clarify the timing and frequency of symptoms measurements and vaccination history

Response: We added how and when the information about symptoms, temperature, and vaccination history were collected. Also, we added a statement about the timing of symptoms and vaccination, as follows:

All of the participants were physically examined and historical data, including age, sex, vaccination status, temperature and symptoms (sore throat, arthralgia and/or myalgia, headache, chills, cough and/or throat phlegm, and nasal discharge) were recorded, as was the time to test from symptom onset. Vaccination status indicated whether an influenza vaccine had been administered during that season before symptom onset. The temperature was taken on presentation by an outpatient physician. Symptoms recorded were those participants reported on presentation.

Please clarify the activity of influenza during the time when influenza testing was evaluated. It would be useful to describe whether testing was done when there was high or low activity, as this has implications for the interpretation of test results.

Response: The graph below shows the epidemic status in 2010-2011 in Japan; we can therefore say that the period from December (week 48) 2010 to April 2011 (week 13) was the 'influenza epidemic season'. http://idsc.nih.go.jp/iasr/prompt/graph/sinin1e-1.gif

Please describe how personnel were trained for sample collection and for reading of results from the Verigene system. Sometimes tests that rely on visual interpretation of findings can provide discordant readings and assessments of reliability are useful.

Response: We described who collected samples and carried out the Verigene test. We also added information about training for performing the Verigene procedure, as follows:

The result for each virus type, 'Detected' or 'Not detected,' was displayed on a monitor. Approximately 2.5 hours was required from sample procurement to final readout.
Specimen collection was performed by one of 5 physicians, who also read the results. These physicians were trained by an instructor from the manufacturer before the study commenced.

Please clarify if the integrity of the collected samples was evaluated – several studies use a house-keeping gen (e.g. RNP) to verify this

Response: We did not evaluate the samples with any other methods.

Please clarify if the physician attending the patient decided on the use of the Verigene referent (page 8)?– this is very important as selection of certain patients may affect the estimation of tests characteristics (e.g. selection of sicker patients for testing)

Response: There was no discretion exercised on the part of the attending physician concerning participant selection. As stated in our answer to a previous query (above), the inclusion criteria are described more fully in the revised manuscript, and the enrolled patients were consecutive. Our hospitals are located in central Tokyo and most patients are young or middle-aged healthy workers, so the participants were relatively young and healthy. We revised the manuscript further, as follows:

RESULT
A total of 82 consecutive patients meeting eligibility criteria were enrolled from December 2010 to April 2011 (Juntendo University Hospital: 37 patients; Juntendo University Nerima Hospital: 45 patients). There was no selection discretion on the part of the attending physicians.
Limitation
The 82 subjects enrolled in the current study were relatively young and healthy. This is because our hospitals are located in central Tokyo and most patients are young or middle-aged healthy workers. We did not exclude elderly patients or patients with co-morbid diseases.

It is unclear why there were only 82 patients in the sample? Were tests done in subjects not included in the 82 described in the report, if so how they compare with those included in the report? How were these patients selected? Was this number considered enough to support the diagnostic evaluation of RIADT (i.e. provided enough statistical power)?

Response: The inclusion criteria are described above. We did not exclude sicker patients or older patients. We enrolled only 82 patients for this study. There were no other subjects who were tested with both RIADT and Verigene procedures. If we detect a significant difference of symptoms between the FN group and TN group, for α=0.05, power=0.80, the calculated required sample size is: FN group=67 and TN group=67. Our number of participants was not enough to sufficiently power the study. We plan to conduct a future study in which more patients are enrolled.

Results
Please provide 95% confidence intervals for all estimations, so the reader could appreciate the precision of the determinations

Response: We added 95% confidence intervals to the revised manuscript, as follows:

When the Verigene VRV test was used as the gold standard, the RIADT sensitivity, specificity, PPV and NPV were 72.9% (95% CI, 61.5-84.2), 91.3% (79.7-102.8), 95.6% (89.5-101.6) and 56.8% (40.8-72.7), respectively.
Combining the RIADT result and presence of temperature ≧37.8 °C or chills increased the sensitivity and the NPV from 72.9% to 96.6% (95% CI, 92.0-101.2%) and from 56.8% to 90.5% (77.9-103.0%), respectively. The specificity and the PPV were 82.6% (67.1-98.1%) and 93.4% (87.2-99.7%), respectively.

In tables 3 & 4, because of the small sample size and apparent distributions, suggest use non-parametric tests for comparison of continuous variables. Please clarify timing of the measurements

Response: We changed the statistical methods used. We used the Wilcoxon rank sum test for continuous variables, and the Fisher's exact test for categorical variables. (Table 3 & 4)

Continuous variables (age, the time from symptom onset, and temperature) were analyzed by the Wilcoxon rank sum test, and the Fisher's exact test was used for comparing patient sex, vaccination status and symptoms. Significance was assigned to results having P values <0.05, and borderline significance was assigned to P values >0.05 and <0.10.

Discussion

Please provide references to support the statement that PCR for influenza is rarely used clinically in Japan – since antivirals appear to be used frequently, it would be useful to clarify if rapid antigen detection is widely used to decide therapy in Japan

Response: We added a reference about use of PCR for influenza in Japan, as follows:

However, these tests take much time and are costly, and so they are rarely used clinically in Japan.[24]

24 Boku S, Naito T, Murai K, et al. Near point-of-care administration by the attending physician of the rapid influenza antigen detection immunochromatography test and the fully automated respiratory virus nucleic acid test: contribution to patient management. Diagnostic Microbiology and Infectious Disease 2013; 76: 445-9.

Please provide a reference for the values presented in the second paragraph about the performance of RIADT and clarify what was the test used as reference.

Response: We added a reference for RIADT, and its reference standard was viral culture and PCR. The manuscript was revised as follows:

According to our RIADT correlative examination results for type A influenza, the sensitivity is 94.3% and specificity is 97.8%, and the values for type B influenza are 87.0% and 100%, respectively, referencing results of virus culture and PCR.[22]

22 The attachment of RapidTesta FLU II (Sekisui Medical, Tokyo, Japan), http://www.eidia.co.jp/product/diagnose/attach/attach_pdf/ippan/rapid-flu-2.pdf (Accessed on October 30, 2013).

Please clarify the correlation of chills with fever – which could be more objective. Clarifying the timing and frequency of the measurements (e.g. at home when symptoms started, or at the clinic after antipyretics were taken) is very important for a proper interpretation of these findings.

Response: The information about chills (and other symptoms) was collected when an outpatient physician examined the participant. The physician asked him/her if he/she had chills (and other symptoms) 'now', and took his/her temperature. Then the physician filled out the recording document. Only a few participants had used antipyretics and their temperature was around 36°C, but we could not analyze the data with regard to antipyretic use, because we asked about symptoms 'now.'

METHODS

All of the participants were physically examined and historical data, including age, sex, vaccination

status, temperature and symptoms (sore throat, arthralgia and/or myalgia, headache, chills, cough and/or throat phlegm, and nasal discharge) were recorded, as was the time to test from symptom onset. Vaccination status indicated whether an influenza vaccine had been administered during that season before symptom onset. The temperature was taken on presentation by an outpatient physician. Symptoms recorded were those participants reported on presentation. Only a few patients had taken antipyretics and their temperature was around 36°C, but we could not analyze the data with regard to antipyretics use.

When describing sensitivity, specificity, etc. it is useful to describe the test used as reference. Please provide this information for the descriptions of Verigene diagnostic characteristics

Response: We added information about the reference standard used for Verigene diagnostic accuracy, as follows:

When direct fluorescent antibody identification and viral culture were used as the gold standards, the sensitivity of the VRV system used in this study for influenza A was 98.7% (95% CI, 96.8-99.5%) and the specificity was 93.2% (95% CI, 91.1-99.9%). For influenza B, the sensitivity was 100% (95% CI, 91.8-100%) and the specificity was 99.7% (95% CI, 99.1-99.9%).[32]
32 FDA 510(k) Summary of K103209: Verigene Respiratory Virus Plus Nucleic Acid Test on the Verigene System (RV+). http://www.accessdata.fda.gov/cdrh_docs/pdf10/K103209.pdf (Accessed on October 30, 2013)

We think the conclusions should address the issue of limited sensitivity first, as this could have direct implications for interpretation of testing and to decide whether additional tests may be needed.

Response: We added a statement about interpretation of testing and additional tests needed, as follows:

CONCLUSIONS
Consistent with previous reports, the sensitivity of the RIADT used in this study was low, due to early administration of the test. Administration of an RIADT too early after symptom onset causes false negative results. In the influenza epidemic season, practitioners should not use RIADT for patients with upper respiratory symptoms and high fever for at least 12 hours after onset. A positive RIADT result after this gives the physician firm support for a diagnosis of influenza. A negative RIADT result does not mean ‘no influenza’. Presence of high fever and chills might predict influenza, but additional tests are necessary for patients with specific symptoms inconsistent with a diagnosis of influenza virus infection.

I would suggest using the STARD statement for reporting of diagnostic tests characteristics (rather than STROBE)

Response: We used the STARD statement instead of the STROBE statement in the revised manuscript.