Clinical diagnosis of seasonal influenza by physicians: a retrospective observational study

Hiroki Maita, Tadashi Kobayashi, Takashi Akimoto, Fumihiko Matsuoka, Shigeki Funakoshi, Hiroshi Osawa, Hiroyuki Kato

ABSTRACT

Objective To elucidate the diagnostic accuracy of pretest probability of influenza (%) by physicians and the factors affecting the clinical diagnosis.

Design Retrospective, single-centre observational study.

Setting A community primary care clinic in Japan.

Participants The participants were recruited from a database of studies conducted during the influenza season from December 2017 to April 2019.

Primary outcome measure Sensitivity and specificity of the physician's clinical diagnosis of influenza recorded in the medical record as pretest probability.

Results A total of 335 patients (median age, 31 years; male, 66.6%) were analysed in this study. The area under the curve (AUC) of the physician’s pretest probability was 0.77. At a cut-off value of 30%, the sensitivity and negative likelihood ratio were 0.92% (95% CI 86.7 to 95.7) and 0.19 (95% CI 0.11 to 0.33), respectively. At a cut-off value of 60%, the specificity and positive likelihood ratio were 90.8% (95% CI 85.4 to 94.6) and 4.01 (95% CI 2.41 to 6.66), respectively. The AUCs of patients who had and had not taken any medications before visiting the clinic were 0.77 (95% CI 0.69 to 0.85) and 0.78 (95% CI 0.71 to 0.84), respectively. The AUCs of patients with type A and B influenza were 0.78 (95% CI 0.72 to 0.84) and 0.76 (95% CI 0.70 to 0.82), respectively. The AUCs of vaccinated and unvaccinated patients were 0.80 (95% CI 0.72 to 0.88) and 0.76 (95% CI 0.63 to 0.89), respectively. The AUC for patients less than 12 hours after onset was 0.69 (95% CI 0.51 to 0.88), and that for patients aged younger than 6 years was 0.69 (95% CI 0.49 to 0.88).

Conclusions The physician’s pretest probability of influenza (%) may be useful for both definitive and exclusionary diagnoses within the limits of our study.

INTRODUCTION

Seasonal influenza is a common disease estimated to affect one billion individuals worldwide annually. In general, the diagnosis of influenza can be established clinically based on the epidemic situation and information from patients. However, establishing the clinical diagnosis of influenza by physicians has been reported to have low diagnostic accuracy. In Japan, the use of rapid influenza diagnostic tests (RIDTs) is the standard for seasonal influenza diagnosis. Clinical diagnosis can be divided into qualitative diagnosis, which is a binary ‘yes/no’ diagnosis of a disease, and quantitative diagnosis, which uses a continuous variable expressed as a percentage or other measures of the likelihood of a disease. Similar to the clinical diagnosis by physicians, the qualitative diagnostic accuracy of influenza self-diagnosis by patients and influenza diagnosis by guardians for their children was reported to be low. However, quantitative self-diagnosis and guardians’ diagnoses have been reported to be useful for influenza diagnosis. There are no reports analysing the diagnostic accuracy of physicians’ quantitative clinical diagnosis of influenza, such as the pretest probability of influenza (%). Pretest probability is the estimated probability of a disease before the test result is known; it is usually based on the physician’s personal experience, local prevalence data or published reports. Various factors may affect the clinical diagnosis of influenza. It has been reported that elderly patients are less likely to develop fever, and vaccinated patients are less likely to develop severe influenza. Medication...
prior to coming to the hospital can also affect the patient’s symptoms. In these cases, the characteristic symptoms of influenza are masked, which may reduce the accuracy of clinical diagnosis by physicians. However, there are no reports that have analysed whether these factors affect the clinical diagnostic accuracy of the physicians.

By analysing the physicians’ clinical diagnosis of influenza and the factors that affect the accuracy of the diagnosis, it is possible to distinguish between cases where the physician’s clinical diagnosis is reliable and cases where further testing is necessary. As a result, we may be able to avoid unnecessary tests for influenza diagnosis, which may cause physical and financial burdens on patients and infection risks to healthcare workers. This study aimed to elucidate the diagnostic accuracy of physicians’ pretest probability of influenza determination (%) as a quantitative clinical diagnosis of seasonal influenza and the factors that affect the clinical diagnosis.

**METHODS**

We conducted a retrospective observational study to analyse the accuracy of the clinical diagnosis of influenza by physicians together with factors affecting the clinical diagnosis using data obtained from previous prospective observational studies conducted at a community clinic (Rokkasho Centre for Community and Family Medicine) and information from the patients’ electronic medical records.

**Patients**

The inclusion criteria for this study were as follows: (1) patients who had participated in a previous study regarding the accuracy of self-diagnosis of seasonal influenza in primary care medical institutions in Japan (Hirosaki University Graduate School of Medicine Ethics Committee, approval number 2017-1100) and (2) patients whose pretest probability of influenza by physicians could be extracted from their medical records. The exclusion criteria were as follows: (1) patients who had not provided consent and (2) patients with missing data.

The previous studies, including the subject of this study, were prospective observational studies conducted during the influenza season from December 2017 to April 2019. All patients with suspected influenza prior to physician consultation completed a pre-examination checklist, and diagnostic accuracy was investigated for self-diagnosis in patients ≥12 years old and guardian’s diagnosis in patients <11 years, using RIDTs as the reference standard. Statistical analyses

**Patient and public involvement**

There was no public or patient involvement in the design, conduct or presentation of the results of the study.

**Physicians**

The clinical staff of the research clinic consisted of two staff physicians (family physicians with more than 20 years of experience), two family medicine residents and other part-time physicians. These physicians worked independently. The physician’s clinical influenza diagnosis, estimated as a pretest probability (recommended, not essential), was immediately recorded as part of the medical record information through medical interviews and examinations. RIDT was then ordered, and medication was prescribed if needed after the results were confirmed. The clinic also functioned as an educational centre for family medicine residents, and each physician and resident recorded pretest probabilities as far as possible for clinical discussion and resident training. The review meetings regarding the patients were held at the end of the daily practice; therefore, the recorded pretest probabilities were not influenced by other physicians.

**Data collection**

We collected data to investigate the diagnostic accuracy of the physicians’ clinical diagnoses of influenza and conduct an exploratory investigation of factors affecting the diagnostic accuracy.

The data extracted from the database of past observational studies were as follows: age, sex, medical history of influenza infection, influenza vaccination status, whether medication had been taken prior to medical visits, clinical signs (axillary temperature at the clinic, axillary temperature at home, pulse rate at the clinic), clinical symptoms (headache, nasal discharge, cough, joint and muscle pain, fatigue, history of fever (acute or sudden, gradual), time of symptom onset, the severity of current symptoms (compared with having a common cold)), time of RIDT, results of RIDT, pretest probability of influenza by the physician and the physician’s final diagnosis.

The data extracted from the medical records were as follows: type of physician (staff physician, resident, others) and the pretest probability of influenza determined by the physician.

**Diagnosis of influenza**

An RIDT (Prime Check Flu, Alfresa, Tokyo, Japan) was used as the reference standard for influenza, and the results were assessed by a clinical laboratory technician independent of the examining physician and the nurse who performed the pre-examination.

**Statistical analyses**

Receiver operating characteristic (ROC) curve analysis was performed to estimate the optimal cut-off point and area under the curve (AUC), which was determined to evaluate the discriminatory power of the physician’s diagnosis under various conditions. Moreover, 2×2 tables were analysed to calculate the sensitivity (Sn), specificity (Sp) and likelihood ratio at each cut-off point.

Based on previous studies, we estimated the AUC to be 0.75 and the influenza-positive rate to be 50%, with an alpha value of 0.05 and a power of 0.9. The sample size required for AUC analysis was estimated to be a minimum of 42 subjects and for subgroup analysis to be at least 100 subjects. All statistical analyses were performed using EZR.
RESULTS

Of the 504 patients who had participated in the previous studies, we analysed 335 patients (median age (IQR 25% to 75%), 31 (11 to 45); male, 66.6%) for whom the physician’s pretest probability could be extracted from the electronic medical records (online supplemental table S1, figure 1). First, the diagnostic accuracy of the physician’s pretest probability was analysed using the ROC curve. The AUC of the physician’s pretest probability was 0.77 (95% CI 0.72 to 0.82) (figure 2). The optimal cut-off value was 50%, for which the Sn and Sp were 79.0% (95% CI 71.9% to 85.0%) and 65.3% (95% CI 57.7% to 72.4%), respectively. At a cut-off value of 30%, the Sn and negative likelihood ratio were 92.0% (95% CI 86.7% to 95.7%) and 0.19 (95% CI 0.11 to 0.33), respectively. At a cut-off value of 80%, the Sp and positive likelihood ratio were 90.8% (95% CI 85.4% to 94.6%) and 4.01 (95% CI 2.41 to 6.66), respectively. The Sn, Sp and likelihood ratio for each of the other cut-off values are listed in table 1.

Among the patients included in the study, 86 (25.7%) had a pretest probability of less than 30%, and 76 (22.7%) had a physician’s pretest probability of 80% or more.

Next, we analysed the diagnostic accuracy of the physicians in each subgroup. The AUCs of patients who had and had not taken some medications before visiting the clinic were 0.77 (95% CI 0.69 to 0.85) and 0.78 (95% CI 0.71 to 0.84), respectively, showing no evident difference. The AUCs of patients with type A and B influenza were 0.78 (95% CI 0.72 to 0.84) and 0.76 (95% CI 0.70 to 0.82), respectively, showing no evident difference. The AUCs of vaccinated and unvaccinated patients were 0.80 (95% CI 0.72 to 0.88) and 0.76 (95% CI 0.63 to 0.89), respectively, showing a slightly higher trend in the vaccinated group. The AUC of family medicine residents was 0.83 (95% CI 0.74 to 0.91), which was higher than that of staff physicians and other physicians. The AUC for patients less than 12 hours after onset was 0.69 (95% CI 0.51 to 0.88), and the AUC for patients aged younger than 6 years was 0.69 (95% CI 0.49 to 0.88), which was the lowest among the age groups (online supplemental table S2).

DISCUSSION

The optimal cut-off value for the physician’s pretest probability (%) was 50%, and the Sn and Sp were 79.0% and 65.3%, respectively. At a cut-off value of 30%, the Sn and negative likelihood ratio were 92.0% and 0.19, respectively, indicating that exclusion diagnosis by clinical diagnosis was useful. At a cut-off value of 80%, the Sp and positive likelihood ratio were 90.8% and 4.01, respectively, indicating that it was useful to establish a definitive diagnosis by clinical diagnosis. Neither medication taken prior to the patient’s visit nor the type of influenza affected the physician’s clinical diagnosis, and vaccination did not make the clinical diagnosis difficult. However, clinical diagnosis is difficult in patients with symptoms less than 12 hours after onset and in paediatric patients aged younger than 6 years.

The usage of the RIDT in Japan and the accuracy of RIDT as a reference standard

RIDT kits are used frequently in Japan. According to Japanese government data, 22.7 million influenza rapid diagnostic kits were supplied to medical institutions during the 2019/2020 influenza season. In many countries, it seems normal not to perform RIDT when the physician’s prior probability is high or low, whereas, in Japan, the threshold for performing RIDT is low for the following reasons. (1) Schoolchildren with influenza are prohibited from attending school until they can no longer transmit the infection to others, according to the School Health and Safety Act. (2) Many companies have in-company rules based on the School Health and Safety Act. Schools and companies usually require patients to submit
Table 1  Diagnostic accuracy of physician's clinical diagnosis of influenza

<table>
<thead>
<tr>
<th>Cut-off value (%)</th>
<th>Sn (% (95% CI))</th>
<th>Sp (% (95% CI))</th>
<th>LR+ (95% CI)</th>
<th>LR− (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>98.8 (95.6 to 99.9)</td>
<td>8.1 (4.5 to 13.2)</td>
<td>1.08 (1.03 to 1.13)</td>
<td>0.15 (0.04 to 0.66)</td>
</tr>
<tr>
<td>20</td>
<td>93.2 (88.2 to 96.6)</td>
<td>30.6 (23.9 to 38.1)</td>
<td>1.34 (1.21 to 1.50)</td>
<td>0.22 (0.12 to 0.41)</td>
</tr>
<tr>
<td>30</td>
<td>92.0 (86.7 to 95.7)</td>
<td>42.2 (34.7 to 49.9)</td>
<td>1.59 (1.39 to 1.82)</td>
<td>0.19 (0.11 to 0.33)</td>
</tr>
<tr>
<td>40</td>
<td>80.9 (74.0 to 86.6)</td>
<td>62.4 (54.8 to 69.7)</td>
<td>2.15 (1.75 to 2.65)</td>
<td>0.31 (0.22 to 0.43)</td>
</tr>
<tr>
<td>50</td>
<td>79.0 (71.9 to 85.0)</td>
<td>65.3 (57.7 to 72.4)</td>
<td>2.28 (1.83 to 2.84)</td>
<td>0.32 (0.23 to 0.44)</td>
</tr>
<tr>
<td>60</td>
<td>64.2 (56.3 to 71.6)</td>
<td>75.7 (68.6 to 81.9)</td>
<td>2.64 (1.98 to 3.52)</td>
<td>0.47 (0.38 to 0.59)</td>
</tr>
<tr>
<td>70</td>
<td>52.5 (44.5 to 60.4)</td>
<td>82.1 (75.5 to 87.5)</td>
<td>2.93 (2.06 to 4.16)</td>
<td>0.58 (0.49 to 0.69)</td>
</tr>
<tr>
<td>80</td>
<td>37.0 (29.6 to 45.0)</td>
<td>90.8 (85.4 to 94.6)</td>
<td>4.01 (2.41 to 6.66)</td>
<td>0.69 (0.61 to 0.79)</td>
</tr>
<tr>
<td>90</td>
<td>14.8 (9.7 to 21.2)</td>
<td>97.7 (94.2 to 99.4)</td>
<td>6.41 (2.27 to 18.07)</td>
<td>0.87 (0.82 to 0.93)</td>
</tr>
</tbody>
</table>

LR−, negative likelihood ratio; LR+, positive likelihood ratio; Sn, sensitivity; Sp, specificity.
sufficient evidence at the time. When dealing with multiple diseases of different severities at the same time, it is necessary to carefully consider the priority of testing. In particular, when preventing droplet-transmitted infections, it is reasonable to avoid testing with nasal wipes and pharyngeal wipes as much as possible. Additionally, with limited medical resources such as medical expenses and medical staff, expensive and manpower-intensive tests should be performed only when they are clearly useful. Further validation studies are needed; however, if RIDT could be reduced by half in Japan, approximately 10 million RIDTs annually could be replaced by clinical diagnosis.19

It is reported that physicians overestimate the pretest probability for common diseases. One possible reason for this could be that physicians often do not think in terms of probability.28 Estimates of pretest probabilities generally reflect clinical knowledge and experience, but can also be derived from epidemiological data. When influenza is circulating in the community, patients with both cough and fever within 48 hours of symptom onset are reported to have a 79% pretest probability of influenza.2 We believe that diagnosing diseases with quantitative indexes such as pretest probability will lead to a more accurate and rational clinical diagnosis.

Limitations
This study has four limitations. First, this was a retrospective study in which new medical record data were added to the database of previous studies. The limited number of influenza cases for which we were able to extract the physician pretest probabilities was included in the study, which may have been affected by selection bias. Larger prospective studies that set physicians’ clinical diagnostic accuracy as the primary outcome are needed. Second, RIDT was used as the reference standard for influenza in this study. Although PCR should be used as a reference standard as it has higher Sn and Sp, it is difficult to perform in primary care settings in Japan. It is necessary to validate these results using reference standards with higher diagnostic accuracy in the future. Third, the number of doctors included in the study was small. The study was conducted in a single community medical institution, making it difficult to generalise the results. Since the results are expected to change depending on the background of the physicians (specialty, years of experience and region of work), it is necessary to conduct a study with a larger number of physicians and physicians with various backgrounds. In addition, the accuracy of the physician's clinical diagnosis needs to be verified in other subjects and settings (nursing home residents, healthcare workers). In particular, this study included only a small number of elderly people, making it difficult to generalise the results to this population. Fourth, the results of the study may not be adaptable in making it difficult to generalise the results to this population. This study included only a small number of elderly people (nursing home residents, healthcare workers). In particular, a larger number of physicians and physicians with various backgrounds, including physicians with various medical specialties, should be included in the study, making it difficult to generalise the results. Since the results are expected to change depending on the background of the physicians (specialty, years of experience and region of work), it is necessary to conduct a study with a larger number of physicians and physicians with various backgrounds. In addition, the accuracy of the physician’s clinical diagnosis needs to be verified in other subjects and settings (nursing home residents, healthcare workers). In particular, this study included only a small number of elderly people, making it difficult to generalise the results to this population.

CONCLUSION
The pretest probability of influenza (%) as a physician’s quantitative clinical diagnosis is useful for definitive and exclusionary diagnoses within the limits of our study. Neither the medication taken before the visit, influenza type nor vaccination affected the physician’s clinical diagnosis. However, clinical diagnosis was difficult for patients with influenza less than 12 hours after onset and children aged younger than 6 years as well as for older age groups.

Acknowledgements The authors would like to thank Shunsuke Soma, MD, and Risa Yamauchi, MD, for their clinical support. We also thank the staff of the Rokkasho Center for Community and Family Medicine and all the patients who participated in our study.

Contributors HM conceived the idea and wrote the original draft of the manuscript. HM, TK, TA and FM were responsible for data acquisition and analysis. TK, TA, FM and SF developed the theory, and HO and HK supervised the findings of this study. All authors discussed the data and commented on the manuscript. HM, TK and TA revised and edited the manuscript. All authors have approved the final manuscript before submission. HM is responsible for the overall content as guarantor.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

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

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The data analysed in the current study were mostly included in this article. Additional data are available from the corresponding author on reasonable request.

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 Hiroki Maita http://orcid.org/0000-0001-9642-9116

REFERENCES


