International travels and fever screening during epidemics: a literature review on the effectiveness and potential use of non-contact infrared thermometers

D Bitar (d.bitar@invs.sante.fr)[1], A Goubar[1], J C Desenclos[1]
1. Department of infectious diseases, Institut de Veille Sanitaire, Saint Maurice, France

Several countries plan to introduce non-contact infrared thermometers (NCIT) at international airports in order to detect febrile passengers, thus to delay the introduction of a novel influenza strain. We reviewed the existing studies on fever screening by NCIT to estimate their efficacy under the hypothesis of pandemic influenza. Three Severe Acute Respiratory Syndrome (SARS) or dengue fever interventions in airports were excluded because of insufficient information. Six fever screening studies in other gathering areas, mainly hospitals, were included (N= 176 to 72,327 persons; fever prevalence= 1.2% to 16.9%). Sensitivity varied from 4.0% to 89.6%, specificity from 75.4% to 99.6%, positive predictive value (PPV) from 0.9% to 76.0% and negative predictive value (NPV) from 86.1% to 99.7%. When we fixed fever prevalence at 1% in all studies to allow comparisons, the derived PPV varied from 3.5% to 65.4% and NPV was =>99%. The low PPV suggests limited efficacy of NCIT to detect symptomatic passengers at the early stages of a pandemic influenza, when fever prevalence among passengers would be =<1%. External factors can also impair the screening strategy: passengers can hide their symptoms or cross borders before symptoms occur. These limits should be considered when setting up border control measures to delay the pandemic progression.

Introduction

The emergence of Severe Acute Respiratory Syndrome (SARS) in 2003 underlined the role of international travels in the rapid spread of infectious diseases and prompted countries to set up border control strategies, in order to reduce the risk of introduction of an infection. Traditional measures such as information for travellers, self-completion of health cards or visual inspection of passengers were implemented by most countries. In addition, non-contact infrared thermometers (NCIT) were introduced in some international airports or other gathering areas such as bus or railway stations. The principle of NCIT is that heat emitted by any organic body can be detected in the infrared radiation spectrum through a remote sensor and transformed into colour images on a monitor. The clinical applications of non-contact infrared thermography include the diagnostic of inflammatory disorders or cancers, or the surveillance of body temperature in neonatology wards through the monitoring of changes in the skin perfusion over time [1,2]. Yet, the extension of these non-invasive diagnostic tools to a public health application, for instance mass screening of breast cancers or fever, has not been thoroughly assessed. Early reports from the SARS experience suggested a low efficacy of NCIT at international airports [3-5] and some authors stressed that NCIT were not currently manufactured for a fever screening purpose [6]. Nevertheless, given the increasing threat of pandemic influenza, some countries envisage to introduce thermal screening at their borders [7]. Their objective is to delay the introduction of the infection through the early detection and isolation of the first infected cases, thus allowing for more time to organise the response.

In this review we summarise the available information on the sensitivity, specificity and predictive values of NCIT used with the objective of fever screening, in airports or other gathering areas. We discuss their potential benefits under the hypothesis of pandemic influenza.

Materials and methods

We performed a systematic MEDLINE search on the literature from 1975 to August 2008. We used the following key words: fever; screening; non-contact, infrared thermography or thermometers; thermal imagers or scanners or pyrometers; thermal screening. The apparent redundancy for some words was necessary because there did not seem to be a standardised vocabulary for the subject. Among the abstracts identified through these key words, we selected the publications which provided the sensitivity and specificity values of NCIT used in a fever screening objective, whatever the cause of the fever.

For international airports, we found partial data from three mass screening interventions using NCIT: two aimed at detecting SARS among international passengers in Canada [4] and in Singapore [8] and one aimed at detecting dengue fever in Taiwan [9]. The numbers of passengers screened and those subsequently confirmed as SARS or dengue cases were provided in these publications but the numbers of passengers who presented with fever due to another cause, i.e. the total numbers of true positive cases, were not available. We therefore discarded these publications which did not allow to derive the sensitivity, specificity and predictive values of NCIT to screen a fever of any origin.

Check for updates

Our search also focused on fever screening interventions in other settings than airports. We selected those which were carried out under conditions considered to be close to a mass screening at international airports, for instance studies implemented in gathering areas, with no preliminary selection or preparation of the tested subjects. We found six studies, performed mainly in hospitals, in which all subjects who were present and accepted to participate were tested. These selected studies summarised in Table 1, included: one in Singapore [10], two in Hong Kong [11,12], two in Taiwan [13,14] and finally one in France [15]. In all, temperatures measured by NCIT were compared to reference values measured by tympanic thermometers i.e. contact thermometers. The authors considered that tympanic thermometers reflected the actual core body temperature with enough confidence, were easy to use because they were routinely used in many hospitals and were more acceptable for the tested subjects than rectal thermometers. The positive and negative predictive values (PPV and NPV) were reported in three of the selected studies. For the others, we derived these values from the available sensitivity, specificity and prevalence data. Finally, because the prevalence of fever in the study populations varied and in order to allow comparisons, we assumed a fixed fever prevalence of 1% in all studies and derived predictive values based upon the sensitivity and specificity as reported in each study. We considered that 1% prevalence was a plausible assumption of the proportion of febrile subjects among international passengers, based on findings from a review of interventions to control SARS [3].

Through our search, we also identified other studies on NCIT with sensitivity and specificity values but these were discarded because they were carried out under strict surrounding conditions which did not fit with our specific objective which was to assess the performances of NCIT under mass screening conditions, in crowded/gathering areas. For instance, participating subjects were asked to refrain from drinking caffeine-based beverages or from exercising the day before. Elsewhere, the device was scanned across the forehead in order to identify specific skin areas where the physiological variations of the skin temperature were reduced. Finally, we also excluded a large number of reports identified through Internet searches, other than Medline, in which information was too scarce.

Results

The study populations ranged from 176 to 72,327 persons (Table 1). They were composed of either hospitalised patients, or persons presenting for emergency or for outpatient consultations, or supposedly healthy persons selected among hospital visitors or sports clubs. Information on age or gender was mostly unavailable. The fever thresholds varied between 37.5°C and 38°C (these were mainly based on the thresholds which were used in the respective countries during the SARS outbreak). The body areas targeted by NCIT systematically included the forehead; the inner eye corner or the external auricular meatus were other skin areas occasionally targeted by the devices. Different types of devices were tested. In four studies, hand-held thermometers were assessed. This implied a shorter distance between the device and tested subjects (=<50cm) than in the two other studies which used remote sensors linked to a monitor (>=50cm). The devices were calibrated according to the respective producers' recommendations. Two studies were carried out in stable external environments consisting of a single dedicated room with stable ambient temperature and ventilation system [12,14].

The prevalence of fever measured by reference contact tympanic thermometers varied from 1.2% to 20.7% in the respective samples, with variable fever thresholds (Table 2). This prevalence was either based on the entire study population or was estimated from a sub-sample. The sensitivity, specificity and predictive values of NCIT targeting the forehead area largely differed between studies. The sensitivity varied from 4.0% to 89.6%, the specificity from 75.4% to 99.6%, the PPV from 0.9% to 76.0% and the NPV from 86.1% to 99.7%. The lowest PPV was found in the study by Chiu

Table 1

Summary of studies on fever screening by non-contact infrared thermometers, 2004-2008

First author, reference	Country, area	Study population (N)	Settings	Sample size *	Temperature threshold	Target area(s)	Device	Environmental conditions
Ng E [10] *Microvasc Res 2004*	Singapore	502	Hospital	310	37.7° C	Forehead	Flir ® S60 Hand held	na
				310		Inner eye corner		
Liu CC [14] *Infect Control Hosp Epidemiol 2004*	Taiwan	500	Outpatient consultation	500	37.5° C	Forehead	Thermofocus ® Hand held	Stable
				500		Auricular meatus		
Chan LS [11] *Travel Med 2004*	Hong Kong	176	Hospital, consultations and sports club	188	37.5°C & 38°C	Forehead	Flir ® -3 models Remote sensors	na
				116		Auricular meatus		
Ng DK [12] *Ann Trop Paed 2005*	Hong Kong	500	Inpatients (Age:1 month-18 years)	500	38° C	Forehead	Standard ST ® Hand held	Stable
Chiu WT [13] *Asia Pac J Public Health 2005*	Taiwan	993	Hospital visitors	993	37.5° C	Forehead	Telesis ® Remote sensors	na
		72.327	Patients + visitors	72.327	37.5° C	Forehead		na
Hausfater P [15] *Emerg Inf Dis 2008*	France	2026	Emergency department (Age 6 – 103 years)	2.026	38° C	Forehead	Rayngen ® ** Hand held	Dedicated nurse

* Number of measurements done in each population
na: Information not available

Check for updates

et al. [13] in their second series of measures conducted among 72,327 patients and hospital visitors, in which fever prevalence was not given.

Receiver operating characteristic (ROC) curves were assessed by three teams; the values of the area under the curve reached 0.96, 0.92 and 0.86 in the studies of E. Ng et al. [10], Hausfater et al. [15] and D. Ng et al. [12], respectively. The correlation coefficient between the forehead and reference tympanic temperatures varied from 0.25 to 0.51 in the two studies where it was quantified [11,14] and was 0.71 when we derived it from the available data in E. Ng [10].

The external auricular meatus area was tested in two studies. This target area yielded higher sensitivity results than the forehead: 82.7% vs. 17.3% [14] and 67.0% vs. 4.0% [11], respectively. Specificity remained high: 98.7% and 96.0%, respectively.

When we fixed the fever prevalence at 1% in all studies and used the sensitivity and specificity values as reported by the respective authors, the derived PPV for the forehead area varied from 3.5% to 65.4% and the derived NPV was =>99% (Table 3).

Discussion

Interpretation and comparison of findings were made difficult by the limited number of selected studies and their wide heterogeneity in terms of methods, study design and environmental conditions. Also, the level of available details in the published papers varied regarding the different study populations which included either healthy or sick persons, and the different types of tested NCIT which included hand-held or remote sensors. The relevance of tympanic (contact) thermometers as reference measurements might also be discussed, but the authors selected feasible and acceptable methods. Another important bias resides in the devices themselves: under operational conditions, the detection of fever by NCIT can be affected by three types of factors [10]. Individual factors such as the consumption of hot beverages or alcohol, pregnancy, menstrual period or hormonal treatments can increase the external skin temperature. Inversely, intense perspiration or heavy face make-up can have a cooling effect on the cutaneous temperature without a parallel decrease of the actual body temperature. The targeted body area scanned by the detector also plays a role, because of physiological differences in vascularisation and consequently in heat distribution. The forehead is subject to important physiological variations but is preferred in screening programmes for feasibility reasons. Inversely, the inner eye corner or the auricular area are less subject to variations but are less accessible: targeting the external auricular area yields better results but travellers would have to be asked to remove their scarves, etc. from around the ear, generating a longer preparation time. Finally, environmental factors can also affect the measurements [2,10], such as the subject-sensor distance, the ambient temperature or humidity and the surrounding ventilation systems, as well as the fact that the person tested should remain immobile for a few seconds in front of the detector.

Despite these constraints, there are several advantages in using NCIT to screen fever at international airports. NCIT save time (temperature is displayed within a few seconds) and reduce close contacts with infected individuals. But, although NCIT appear suitable for entry screening because of high specificity and NPV, the low sensitivity values reported in the studies suggest that the risk of missing febrile individuals (1-sensitivity) would reach 83 to 85% [11,14]. In addition, given the low PPV, hostile reactions may arise among a high proportion of passengers mistakenly classified as febrile by the sensors and subsequently referred for medical examination. Because of these limitations, most authors were extremely cautious in their respective conclusions, stating for instance that NCIT may serve as a proxy tool [11] or that surveillance and contact tracing would be more beneficial [14].

Table 2

Fever screening by non-contact infrared thermometers, 2004-2008: sensitivity, specificity and predictive values according to the body area targeted

First author, country, publication year	Sample size	Target area(s)	Temperature threshold	Fever prevalence %	Sensitivity %	Specificity %	PPV %	NPV %
Ng E Singapore 2004	310	Forehead	37.7°C	16.9	89.6	94.3	***76.0****	***97.8****
	310	Inner eye corner	37.7°C	16.9	85.4	95	***77.7****	***97.0****
Liu CC Taiwan 2004	500	Forehead	37.5°C	-	17.3	98.2		
	500	Auricular meatus	37.5°C	-	82.7	98.7		
Chan LS Hong Kong 2004	188	Forehead	38°C	14.3	4	99	***40.1****	***86.1****
	-	Forehead	37.5°C	Na	15	98		
	116	Auricular meatus	38°C	20.7	67	96	***81.4****	***91.8****
Ng DK Hong Kong 2005	500	Forehead	37.5°C †	12.3 †	89.4	75.4	33.7	98.1
Chiu W Taiwan 2005	993	Forehead	37.5°C	1.2	75	99.6	***69.9****	***99.7****
	72.327	Forehead	37.5°C	-	-	-	***0.9****	
Hausfater P France 2008	2.026	Forehead	38.0°C	3.0	82	77	10	99

* Values derived from the available information are in ***bold italic***
† The 37.5°C cut-off corresponds to the optimal sensitivity and specificity values reported by the authors whereas the prevalence (12.3%) is based on a 38°C threshold.
PPV: Positive predictive values; NPV: Negative predictive values

Check for updates

Under a pandemic influenza scenario, one could expect a higher PPV, because of a higher prevalence of fever (>1%). But it is in the very early stages of the pandemic that NCITs would be considered as a way to delay the introduction of infection in a given area. In these early stages, the number of infected cases would be very low and the overall prevalence of fever among international passengers would remain below the 1% rate which we set in our analysis.

Finally, even if better-performing devices were manufactured and implementation costs were affordable for national authorities, the overall efficiency of the screening intervention would still need to be examined. As stated by an international experts committee [16], the overall sensitivity of border control is likely to be limited. Modelling works show that border control strategies aimed at reducing the risk of introduction of SARS or influenza in a country have poor sensitivity [17] and limited impact [18-21]. The epidemiological characteristics of the infection play a major role, as illustrated by the differences between SARS and influenza. For SARS, infectiousness peaks after the onset of symptoms, therefore early detection of patients may indeed contribute to their early isolation and thus reduce transmission. For pandemic influenza, because it is assumed that infectiousness starts a few hours before the onset of symptoms, some cases would be missed and would generate secondary cases after their entry in the country. Sociological factors can also affect the efficacy of border control measures. Knowing that thermal screening is organised in international airports may motivate some symptomatic passengers to delay their travel, but inversely, others may try to hide their symptoms or by-pass border control [22;23]. The psychological reassuring effect on the public can influence the decision to implement such screening, as was the case in Singapore and Canada [24-26], but these countries also recognised that the public may loose confidence in this measure if an undetected case had entered the country and generated secondary cases. Because public perceptions are important, policy makers may feel some pressure to use NCIT but the decision making process should not ignore the poor scientific evidence on NCIT's efficacy to delay the introduction of a novel influenza strain. For transparency reasons, the surrounding sociological, demographic, epidemiological and environmental factors which can influence the screening strategy must also be taken into consideration.

Table 3

Fever screening by non-contact infrared thermometers (NCIT), 2004-2008: positive and negative predictive values of NCIT for forehead temperature screening, assuming a fever prevalence of 1%

First author, country	Sample	Fever threshold	Sensitivity %	Specificity %	PPV %	NPV %
Ng E. Singapore	310	37.7°C	89.6	94.3	13.7	99.9
Liu CC Taiwan	500	37.5°C	17.3	98.2	8.8	99.2
Chan LS Hong Kong	188	38°C	4	99	3.9	99.0
	188	37.5°C	15	98	7.0	99.1
Ng DK Hong Kong	1.000	37.5°C	89.4	75.4	3.5	99.9
Chiu W Taiwan	993	37.5°C	75	99.6	65.4	99.7
Hausfater P France	2.026	38°C	82	77	9.9	99.3

PPV: Positive predictive values; NPV: Negative predictive values

Aknowledgements

We thank Isabelle Bonmarin, Didier Che, Dennis Falzon and Veronique Vaillant for their helpful comments.

This work was done as part of SARSControl : Effective and Acceptable Strategies for the Control of SARS and New Emerging Infections in China and Europe, a European Commission project funded within the Sixth Framework Programme, Thematic Priority Scientific Support to policies, Contract no. SP22-CT-2004-003824.

References

1. Clark RP, Stothers JK. Neonatal skin temperature distribution using infra-red colour thermography. J Physiol. 1980;302:323-33.
2. Jiang LJ, Ng EY, Yeo AC, Wu S, Pan F, Yau WY, et al. A perspective on medical infrared imaging. J Med Eng Technol. 2005;29(6):257-67.
3. Bell DM. Public Health Interventions and SARS Spread, 2003. Emerg Infect Dis. 2004;10(11):1900-6.
4. Health Canada. Thermal image scanners to detect fever in airline passengers, Vancouver and Toronto, 2003. Can Commun Dis Rep. 2004;30(19):165-7.
5. an CC. SARS in Singapore--key lessons from an epidemic. Ann Acad Med Singapore. 2006;35(5):345-9.
6. Wong JJ, Wong CY. Non-contact infrared thermal imagers for mass fever screening--state of the art or myth? Hong Kong Med J. 2006;12(3):242-4.
7. Mounier-Jack S, Jas R, Coker R. Progress and shortcomings in European national strategic plans for pandemic influenza. Bull World Health Organ. 2007;85(12):923-9.
8. Wilder-Smith A, Goh KT, Paton NI. Experience of severe acute respiratory syndrome in Singapore: Importation of cases, and defense strategies at the airport. J Travel Med. 2003;10(5):259-62.
9. Shu PY, Chien LJ, Chang SF, Su CL, Kuo YC, Liao TL, et al. Fever screening at airports and imported dengue. Emerg Infect Dis. 2005;11(3):460-2.
10. Ng EY, Kaw GJ, Chang WM. Analysis of IR thermal imager for mass blind fever screening. Microvasc Res. 2004;68(2):104-9.
11. Chan LS, Cheung GT, Lauder IJ, Kumana CR, Lauder IJ. Screening for fever by remote-sensing infrared thermographic camera. J Travel Med. 2004;11(5):273-9.
12. Ng DK, Chan CH, Lee RS, Leung LC. Non-contact infrared thermometry temperature measurement for screening fever in children. Ann Trop Paediatr. 2005;25(4):267-75.
13. Chiu WT, Lin PW, Chiou HY, Lee WS, Lee CN, Yang YY, et al. Infrared thermography to mass-screen suspected SARS patients with fever. Asia Pac J Public Health. 2005;17(1):26-8.
14. Liu CC, Chang RE, Chang WC. Limitations of forehead infrared body temperature detection for fever screening for severe acute respiratory syndrome. Infect Control Hosp Epidemiol. 2004;25(12):1109-11.
15. Hausfater P, Zhao Y, Defrenne S, Bonnet P, Riou B. Cutaneous infrared thermometry for detecting febrile patients. Emerg Infect Dis. 2008;14(8):1255-8.
16. Bell DM. Non-pharmaceutical interventions for pandemic influenza, international measures. Emerg Infect Dis. 2006;12(1):81-7.
17. Samaan G, Patel M, Spencer J, Roberts L. Border screening for SARS in Australia: what has been learnt? Med J Aust. 2004;180(5):220-3.
18. Cooper BS, Pitman RJ, Edmunds WJ, Gay NJ. Delaying the International Spread of Pandemic Influenza. PLoS Med. 2006;3(6):e212.
19. Gumel AB, Ruan S, Day T, Watmough J, Brauer F, van den DP, et al. Modelling strategies for controlling SARS outbreaks. Proc Biol Sci. 2004;271(1554):2223-32.
20. Hollingsworth TD, Ferguson NM, Anderson RM. Will travel restrictions control the international spread of pandemic influenza? Nat Med. 2006;12(5):497-9.
21. Pitman RJ, Cooper BS, Trotter CL, Gay NJ, Edmunds WJ. Entry screening for severe acute respiratory syndrome (SARS) or influenza: policy evaluation. BMJ. 2005;331(7527):1242-3.
22. Desenclos JC, van der WS, Bonmarin I, Levy-Bruhl D, Yazdanpanah Y, Hoen B, et al. Introduction of SARS in France, March-April, 2003. Emerg Infect Dis. 2004;10(2):195-200.
23. Markel H, Gostin LO, Fidler DP. Extensively drug-resistant tuberculosis: an isolation order, public health powers, and a global crisis. JAMA. 2007;298(1):83-6.
24. James L, Shindo N, Cutter J, Ma S, Chew SK. Public health measures implemented during the SARS outbreak in Singapore, 2003. Public Health. 2006;120(1):20-6.
25. St John RK, King A, de Jong D, Bodie-Collins M, Squires SG, Tam TW. Border screening for SARS. Emerg Infect Dis. 2005;11(1):6-10.

Check for updates

26. Svoboda T, Henry B, Shulman L, Kennedy E, Rea E, Ng W, et al. Public health measures to control the spread of the severe acute respiratory syndrome during the outbreak in Toronto. N Engl J Med. 2004;350(23):2352-61.

This article was published on 12 February 2009.

Citation style for this article: Bitar D, Goubar A, Desenclos JC. International travels and fever screening during epidemics: a literature review on the effectiveness and potential use of non-contact infrared thermometers. Euro Surveill. 2009;14(6):pii=19115. Available online: http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19115

Check for updates