DOI:10.1111/j.1750-2659.2011.00198.x
www.influenzajournal.com

Original Article

A cluster randomized clinical trial comparing fit-tested and non-fit-tested N95 respirators to medical masks to prevent respiratory virus infection in health care workers

Chandini Raina MacIntyre,[a] Quanyi Wang,[b] Simon Cauchemez,[c] Holly Seale,[a] Dominic E. Dwyer,[d] Peng Yang,[b] Weixian Shi,[b] Zhanhai Gao,[a] Xinghuo Pang,[b] Yi Zhang,[b] Xiaoli Wang,[b] Wei Duan,[b] Bayzidur Rahman,[a] Neil Ferguson[c]

[a]School of Public Health and Community Medicine, Faculty of Medicine, University of New South Wales, Sydney, Australia. [b]The Beijing Center for Disease Prevention and Control, Beijing, China. [c]MRC Center for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, London, UK. [d]Institute for Clinical Pathology and Medical Research, Westmead Hospital, Sydney, Australia.

Correspondence: Professor C. Raina MacIntyre, Professor of Infectious Diseases Epidemiology and Head of School, School of Public Health and Community Medicine, Faculty of Medicine, University of New South Wales, Samuels Building, Room 325, Sydney, 2052 NSW, Australia.
E-mail: r.macintyre@unsw.edu.au

Accepted 3 December 2010. Published Online 27 January 2011.

Background We compared the efficacy of medical masks, N95 respirators (fit tested and non fit tested), in health care workers (HCWs).

Methods A cluster randomized clinical trial (RCT) of 1441 HCWs in 15 Beijing hospitals was performed during the 2008/2009 winter. Participants wore masks or respirators during the entire work shift for 4 weeks. Outcomes included clinical respiratory illness (CRI), influenza-like illness (ILI), laboratory-confirmed respiratory virus infection and influenza. A convenience no-mask/respirator group of 481 health workers from nine hospitals was compared.

Findings The rates of CRI (3·9% versus 6·7%), ILI (0·3% versus 0·6%), laboratory-confirmed respiratory virus (1·4% versus 2·6%) and influenza (0·3% versus 1%) infection were consistently lower for the N95 group compared to medical masks. By intention-to-treat analysis, when *P* values were adjusted for clustering, non-fit-tested N95 respirators were significantly more protective than medical masks against CRI, but no other outcomes were significant. The rates of all outcomes were higher in the convenience no-mask group compared to the intervention arms. There was no significant difference in outcomes between the N95 arms with and without fit testing. Rates of fit test failure were low. In a *post hoc* analysis adjusted for potential confounders, N95 masks and hospital level were significant, but medical masks, vaccination, handwashing and high-risk procedures were not.

Interpretation Rates of infection in the medical mask group were double that in the N95 group. A benefit of respirators is suggested but would need to be confirmed by a larger trial, as this study may have been underpowered. The finding on fit testing is specific to the type of respirator used in the study and cannot be generalized to other respirators.

Trial registration Australian New Zealand Clinical Trials Registry (ANZCTR), ACTRN: ACTRN12609000257268 (http://www.anzctr.org.au).

Keywords Health workers, influenza, masks, PPE.

Please cite this paper as: MacIntyre *et al.* (2011) A cluster randomized clinical trial comparing fit-tested and non-fit-tested N95 respirators to medical masks to prevent respiratory virus infection in health care workers. Influenza and Other Respiratory Viruses 5(3), 170–179.

Introduction

The current influenza A H1N1 2009 virus pandemic,[1] the ongoing zoonotic transmission of influenza A H5N1 and the emergence of oseltamivir-resistant seasonal influenza A H1N1 are threats to human health. Hospital health care workers (HCWs) are key to effective pandemic response and the capacity of health care systems. Respiratory protection is one of the key non-pharmaceutical interventions for protection of HCWs.

Nosocomial influenza and other outbreaks result in significant morbidity and costs[2,3] and can occur in the absence of community epidemics.[4] During outbreaks of infectious diseases, hospitals may amplify virus transmission, as demonstrated during severe acute respiratory syndrome (SARS).[5] Furthermore, anticipated antiviral shortages and

delays in vaccine development make non-pharmaceutical interventions crucial. There are gaps in knowledge about prevention of influenza by medical masks and respirators. There are several prospective, randomized controlled trials on the use of handwashing,[6–8] but only two trials on the use of medical masks/respirators in households.[9,10] In one of these studies, we showed that medical masks/respirators in compliant users in the household setting were associated with reductions in the risk of influenza-like illness (ILI)-associated infection.[10] To date, there is one small randomized controlled trial (RCT) of medical masks compared to respirators in HCWs[11] which found no difference, but lacked a control arm. Medical masks are not designed to provide respiratory protection.[12] They have consistently lower filtration efficiency when compared to respirators, which are designed specifically for respiratory protection.[13–15] Medical masks were designed to prevent wound contamination when worn by the surgeon; however, three RCTs failed to show efficacy against their intended design.[16–18]

The aim of this study was to determine the efficacy of medical masks compared to fit-tested and non-fit-tested N95 respirators in HCWs in the prevention of disease because of influenza and other respiratory viruses.

Methods

A prospective, cluster randomized trial of medical mask and respirator use in frontline HCWs was conducted from December 2008 to January 2009 in Beijing, China. We initially aimed to determine the efficacy of two different kinds of respiratory protection (N95 respirators and medical masks) during the influenza season compared to each other and compared to a no-mask group. However, although we intended to have a randomized control group, this was not acceptable to the Chinese IRB, who felt it would be unethical to assign HCWs randomly to not wear a mask, given mask use was widespread in Chinese hospitals that were included in the randomization. As such, we studied a convenience-selected no-mask group of HCWs who did not wear a mask. These HCWs were selected from other hospitals where mask wearing was not routine during the study period. Absence of randomization in the no-mask group meant that we eventually had to restrict the primary analysis of the trial to the comparison of the efficacy of N95 respirators and medical masks with each other.

Participants were hospital HCWs aged ≥18 years from the emergency departments and respiratory wards of 15 hospitals. These wards were selected as high-risk settings in which repeated and multiple exposures to respiratory infections are expected. We also monitored all participating wards by active surveillance for clinically compatible illness, including in the no-mask group, for outbreaks of respiratory infection in patients during the study period, and none was detected. All hospitals were large, tertiary hospitals in urban Beijing, and there was no variation in the start of the influenza season within this geographic area.

Recruitment commenced on the 1 December 2008 and final follow-up was completed on 15 January 2009. The study protocol was approved by the Institutional Review Board and Human Research Ethics Committee of the Beijing Ministry for Health. Verbal informed consent was provided by participants, and they were provided written information about the study.

The nine hospitals in the convenience no-arm group were not part of the randomization, but HCWs in those hospitals were selected from the same type of wards as the intervention arms (emergency departments and respiratory wards). They were followed up in the same way as the trial participants for development of infections.

Randomization and intervention

The unit of randomization was hospitals. Hospitals were randomized to one of three intervention arms: (i) Medical masks (3M™ medical mask, catalogue number 1820, St Paul, MN, USA); (ii) N95 fit-tested mask (3M™ flat-fold N95 respirator, catalogue number 9132) and (iii) N95 non-fit-tested mask (3M™ flat-fold N95 respirator, catalogue number 9132). Figure 1 outlines the recruitment and randomization (using a secure computerized randomization program) process. A pre-study assessment of hospital infection control levels determined that the hospitals had sufficient diversity to warrant stratified randomization by size of hospital and level of infection control. This assessment measured ventilation, spatial dimensions, bedding configuration, handwashing facilities and personal protective equipment use. The Ministry of Health in 1989 categorizes hospitals in China into three levels (Level 3 is the highest) depending on their level of sophistication, equipment and staff/bed numbers. Fifteen hospitals were randomized – five level 2 and ten level 3.

Primary endpoints

(i) Clinical respiratory illness (CRI),[19] defined as two or more respiratory or one respiratory symptom and a systemic symptom; (ii) ILI, defined as fever ≥38°C plus one respiratory symptom (i.e. cough, runny nose, etc.); (iii) laboratory-confirmed viral respiratory infection (detection of adenoviruses, human metapneumovirus, coronavirus 229E/NL63, parainfluenza viruses 1, 2 and 3, influenza viruses A and B, respiratory syncytial virus A and B, rhinovirus A/B and coronavirus OC43/HKU1 by multiplex PCR); (iv) laboratory-confirmed influenza A or B and (v) adherence with mask/respirator use.

The choice of a relatively broad CRI definition was dictated by our interest in interrupting transmission of a wide

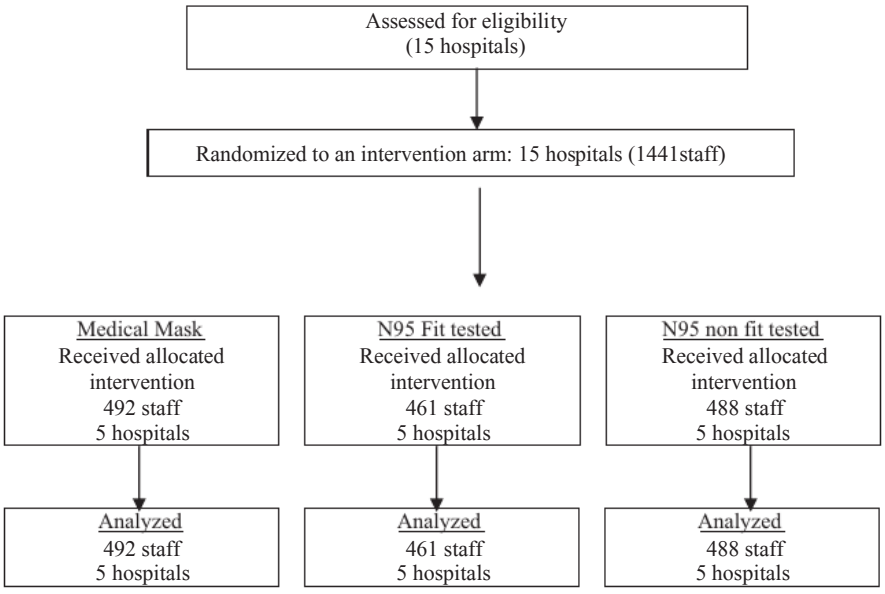

Figure 1. Consort Diagram of recruitment and follow-up.

range of respiratory viruses, which in adults may or may not be accompanied by fever. Also, all respiratory pathogens share a similar transmission mechanism namely aerosol, droplet and fomite spread, although the relative role of these factors may vary between different viruses and in different clinical situations. Other endpoints included adverse effects, measured using a semi-structured questionnaire and adherence.

Eligibility

Any nurse, doctor or ward clerk who worked full time in the emergency or respiratory wards at the hospital were eligible. HCWs were excluded if they: (i) were unable or refused to consent; (ii) had beards, long moustaches or long facial hair stubble; (iii) had a current respiratory illness, rhinitis and/or allergy and (iv) worked part-time or did not work in the aforementioned wards/departments. In all participating wards, 100% of eligible health workers participated.

Intervention

Participants wore the mask or respirator on every shift for 4 consecutive weeks after being shown when to wear it and how to fit it correctly. Participants were supplied daily with either three masks for the medical mask group or two N95 respirators. Participants were asked to store the mask in a paper bag every time they removed it (for toilet breaks, tea/lunch breaks and at the end of every shift) and place the bagged mask or respirator in their locker. All participants were instructed on the importance of hand hygiene prior to/after the removal of medical masks and respirators. Participants in arm two underwent a fit-testing procedure using a 3M™ FT-30 Bitrex Fit Test kit according to the manufacturers' instructions (3M™, St Paul, MN, USA).

Detailed demographic and clinical details of all participants were collected. This included age, sex, smoking history, comorbidities, seasonal influenza vaccination status, medications, conduct of high-risk procedures (defined as suctioning, intubation, nebulized medications, chest physiotherapy and other aerosol generating procedures), handwashing practices, use of other personal protective equipment (gowns, gloves, eye shields and hair/foot covers) and results of laboratory tests. Use of specific interventions for influenza such as antivirals was also measured.

Follow-up

Participants were followed for 4 weeks of wearing the masks or respirators and an extra week of non-wearing for development of respiratory symptoms.

All participants received a mercury thermometer to measure their temperature at the beginning of each day and at the onset of any symptoms. Diary cards were provided for the duration to record daily the (i) number of hours worked; (ii) mask/respirator usage and (iii) recognized CRI encounters.

Participants were contacted daily by phone or face-to-face contact to actively identify incident cases of respiratory infection. At each ward, the head nurse actively followed up all participants and identified incident illness. Staff members from the District CDC also undertook daily monitoring of the sites. If participants were symptomatic, swabs of both tonsils and the posterior pharyngeal wall were collected.

We also monitored adherence with mask or respirator use over the 4-week time course by: (i) observation: the

head ward nurse observed compliance on the ward on a daily basis and recorded the information on a structured form, (ii) self-report: a diary card with tick boxes was given to each subject, to be carried during the day. Adherence to wearing the masks or respirators was monitored by these diary cards and returned to researchers on a weekly basis. Exit interviews with participants were conducted after the 4 weeks to gain further insights into adherence and other issues around the use of masks/respirators including adverse effects.

Sample collection and laboratory testing

Participants with symptoms had two pharyngeal swabs collected by a trained nurse or doctor. Double rayon-tipped, plastic-shafted swabs were used to scratch both tonsilar areas and the posterior pharyngeal wall. These were transported immediately after collection to the laboratory, or at 4°C within 48 hours if transport was delayed.

Pharyngeal swabs were tested with at the Laboratories of the Beijing Centers for Disease Control and Prevention. Viral DNA/RNA was extracted from 300 μl of each respiratory specimen using the Viral Gene-spin^{TM} kit (iNtRON Biotechnology, Inc., Seoul, Korea) according to the manufacturer's instructions. Reverse transcription was performed on 8 μl of RNA in a final reaction volume of 20 μl for 1·5 hours at 37°C, using the RevertAid^{TM} First Strand cDNA Synthesis kit (Fermentas, Burlington, ON, Canada) to synthesize cDNA. Multiplex polymerase chain reaction (PCR) was carried out using the Seeplex® RV12 Detection kit (Seegen, Inc., Seoul, Korea) to detect adenoviruses, human metapneumovirus, coronavirus 229E/NL63, parainfluenza viruses 1, 2 or 3, influenza viruses A or B, respiratory syncytial virus A or B, rhinovirus A/B and coronavirus OC43/HKU1. Three microlitres of synthesized first-strand cDNA, 4 μl of multiplex primers, 10 μl master mix (hot start *Taq* DNA polymerase and dNTP are included in the reaction buffer) and 3 μl of 8-methoxypsoralen (8-MOP) were added (8-MOP, accompanied by UV irradiation for 20 minutes, prevents amplification of contaminated DNA). A mixture of 12 viral clones was used as a positive control template, and sterile deionized water was used as a negative control. After preheating at 95°C for 15 minutes, 40 amplification cycles were carried out under the following conditions in a thermal cycler (GeneAmp PCR system 9700, Foster City, CA, USA): 94°C for 30 seconds, 60°C for 1·5 minutes and 72°C for 1·5 minutes. Amplification was completed at the final extension step at 72°C for 10 minutes. The multiplex PCR products were visualized by electrophoresis on an ethidium bromide-stained 2% agarose gel. Viral isolation by MDCK cell culture was undertaken for some of the influenza samples which were positive by nuclei acid detection. Specimen processing, DNA/RNA extraction, PCR amplification and PCR product analyses were conducted in different rooms to avoid cross-contamination.

Analysis

The primary endpoints of interest as described above were analysed by intention-to-treat analysis. The two N95 arms were also combined and compared to the medical mask arm, given that there was no significant difference between them and rates of fit test failure were extremely low in the fit-tested arm (5/461 fit test failures). Differences in proportions between the trial arms were tested by calculation of Pearson's chi-square using SAS 9.2 software (Cary, NC, USA). The distribution of key potentially confounding variables between study arms was compared. To estimate the odds ratio while adjusting for the clustering effects, we used a random effect logistic regression model. In the model, we added a hospital-specific random intercept in the linear predictors, and maximum likelihood was estimated using adaptive quadrature.[20] The model was fitted using 'xtlogit' command in STATA (College Station, TX, USA).[21]

We also conducted multivariable analysis to adjust for the potential confounders. In the initial model, we included all the variables along with the main exposure variable those were significant ($P < 0{\cdot}05$) in the univariate analysis. We then used a backward elimination method to remove the variables that did not have any confounding effect, that is, could not make meaningful (roughly 10%) change in the effect measure with the main exposure variable.[22] In case of high multi-collinearity because of strong correlation among the potential confounders, we chose the more relevant ones having the highest confounding effect on the association of interest.

We analysed compliance as wearing the mask for >80% of the shift.

Sample size calculation

To obtain 80% power at 2-sided 5% significant level for detecting a significant difference of attack rate between the intervention arms, and for an assumed 5% attack rate in the N95 arm and 12% in the medical mask arm, a sample size of 488 participants or five clusters (hospitals) per arm was required for cluster size (m) 100 and intra-cluster correlation coefficient (ICC) 0·01.[23] The design effect (deff) for this cluster randomization trial was 2 (deff = $1 + (m-1) \times \text{ICC} = 1 + (100-1) \times 0{\cdot}01 = 2$). As such, we aimed to recruit a sample size of 500 per arm.

Results

A total of 1441 nurses and doctors in 15 Beijing hospitals were recruited into the randomized arms and 481 nurses and doctors in nine hospitals were recruited into the convenience no-mask group. Figure 1 shows the recruitment

Table 1. Demographic and other characteristics by arm of randomization

Variable	Medical mask (% and 95% CI) (*n* = 492)	N95 non-fit (% and 95% CI) (*n* = 488)	*P* value	N95 fit (% and 95% CI) (*n* = 461)	*P* value
Sex (male)	57/492 11·6 (8·8–14·4)	40/488 8·2 (5·8–10·6)	0·08	45/461 9·8 (7·1–12·5)	0·36
Age (mean)	32·7 (31·8–33·5)	33 (32·2–33·8)	0·52	35·3 (34·4–36·2)	<0·01
Education*	189/492 38·4 (34·1–42·7)	180/488 36·9 (32·6–41·2)	0·62	155/461 33·6 (29·3–37·9)	0·12
Current smoker	18/492 3·7 (2·0–5·3))	17/488 3·5 (1·9–5·1)	0·88	13/461 2·8 (1·3–4·3)	0·47
Anyone smoking in family	187/492 38·0 (33·7–42·3)	216/488 44·3 (39·9–48·7)	0·05	213/461 46·2 (41·7–50·8)	0·01
Four or more people of family	89/492 18·1 (14·7–21·5)	122/488 25·0 (21·2–28·8)	0·01	113/461 24·5 (20·6–28·4)	0·02
Four or more adults in family	67/492 13·6 (10·6–16·7)	100/488 20·5 (16·9–24·1)	<0·01	90/461 19·5 (15·9–23·1)	0·01
One child in family**	201/485 41·4 (37·1–45·8)	209/464 45·0 (40·5–49·6)	0·26	197/457 43·1 (38·6–47·7)	0·61
No children in family**	278/485 57·3 (52·9–61·7)	244/464 52·6 (48·0–57·1)	0·14	254/457 55·6 (51·0–60·1)	0·59
Influenza vaccination in 2008	109/492 22·2 (18·5–25·8)	105/488 21·5 (17·9–25·2)	0·81	44/461 9·5 (6·9–12·2)	<0·01
Influenza vaccination in 2007	107/492 21·7 (18·1–25·4)	105/488 21·5 (17·9–25·2)	0·93	68/461 14·8 (11·5–18·0)	<0·01
Previous mask wearing					
At work	469/492 95·3 (93·5–97·2)	475/488 97·3 (95·9–98·8)	0·09	431/461 93·5 (91·2–95·7)	0·22
At home	3/492 0·6 (0–1·3)	4/488 0·8 (0–1·6)	0·7	6/461 1·3 (0·3–2·3)	0·27
Public transport	7/492 1·4 (0·4–2·5)	11/488 2·3 (0·9–3·6)	0·33	19/461 4·1 (2·3–5·9)	0·01
Staff (doctors)	154/492 31·3 (27·2–35·4)	144/488 29·5 (25·5–33·6)	0·54	166/461 36·0 (31·6–40·4)	0·12
Undertake handwashing after touching a patient	435/491 88·6 (85·8–91·4)	424/483 87·8 (84·9–90·7)	0·7	382/460 83·0 (79·6–86·5)	0·01
Sick contact in household during trial	13/296 4·4 (2·1–6·7)	21/376 5·6 (3·3–7·9)	0·48	7/283 2·5 (0·7–4·3)	0·21
Hospital level (level 2)	0/492 0	48/488 9·8 (7·2–12·5)	<0·01	319/461 69·2 (65·0–73·4)	<0·01
High-risk procedure	201/492 40·9 (36·5–45·2)	171/488 35·0 (30·8–39·3)	0·06	108/461 23·4 (19·6–27·3)	<0·01

*Undergraduate and above.
**Cases without data are not included.

process. The distribution of demographic variables was generally similar between arms (Table 1), but was significantly different for anyone smoking in the family, four or more people in family, four or more adults in family, influenza vaccination in 2008 and 2007, public transport, handwashing, hospital level and high-risk procedures. In regards to hand hygiene, 83% (382/461), 87·8% (428/488) and 88·6% (435/492) of participants from the N95 fit test arm, N95 non-fit test arm and medical mask arm stated that they washed their hands between patients, respectively.

For all outcomes, non-fit-tested N95 respirators had lower rates of infections compared to fit-tested N95s (for all N95 versus medical masks, the rates were 3·9% versus 6·7% for CRI, 0·3% versus 0·6% for ILI, 1·4% versus 2·6% for laboratory-confirmed virus and 0·3% versus 1% for influenza) but these differences were not significant. All infection outcomes were consistently higher (approximately

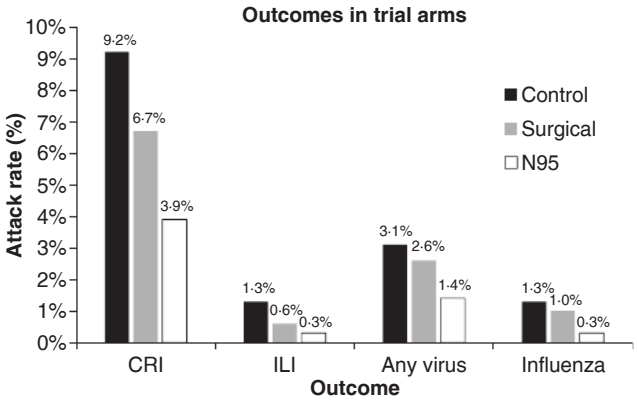

Figure 2. Outcomes in trial arms.

double) in the medical mask group compared to the N95 group (Figure 2). There were no cases of influenza in the non-fit-tested N95 arm, three in the fit-tested N95 arm and five in the medical mask arm. After adjustment for clustering, non-fit-tested N95 masks were significantly protective compared to medical masks against CRI, but other outcomes were not significant between N95 and medical masks (Table 2). When compared to the convenience no-mask group and adjusted for clustering, N95 non-fit-tested was significantly protective against CRI, and all N95 was protective against laboratory-confirmed virus and laboratory-confirmed influenza (Table 3). In a *post hoc* analysis carried out to adjust for potential confounders which were unevenly distributed between arms, all N95 and hospital level remained significant for CRI and laboratory-confirmed viral infection, but handwashing, vaccination and high-risk procedures were not significant (Table 4).

Fit-testing failure rate was very low (5/461, 1·08%). Rates of adherence in all arms of the study were high (Figure 3). Table 5 shows adverse events associated with medical mask or N95 use, and that N95 respirators were associated with higher rates of adverse events. Adherence with mask or respirator wearing was high and not significantly different in all arms, with 74% adherence (95% CI 70–78%) in the N95 fit-tested arm, 68% in the N95 non-fit-tested arm (95% CI 64–73%) and 76% in the medical mask arm (95% CI 72–79%). The duration of mask wearing in these arms, respectively, was 5·2 hours (95% CI 5·1–5·4 hours), 4·9 hours (95% CI 4·8–5·1 hours) and 5 hours (95% CI 4·9–5·2 hours; Figure 3).

Discussion

We found that rates of respiratory tract infection were approximately double in the medical mask group compared to the N95 group in health workers who wore masks throughout their shift. However, only the N95 non-fit-tested arm was significantly protective against CRI, and there were no other significant differences between N95 respirators and medical masks for the four primary outcomes in the adjusted analysis. However, it should be noted that under the null hypothesis where there is no difference between groups, the probability that we wrongly find at least one significant difference given the 12 tests undertaken is 46%. The trial may also be underpowered because observed attack rates were lower than expected.

The rates of all outcomes were higher in the convenience no-mask group than in the masks groups. By adjusted

Table 2. Intention to treat analysis using random effect logistic regression analysis

Arms	CRI		ILI*		Laboratory-confirmed virus**		Influenza	
	N (%)	OR (95% CI)***	*N* (%)	OR (95% CI)***	*N* (%)	OR (95% CI)***	*N* (%)	OR (95% CI)***
N95 fit-tested	21/461 (4·6)	0·76 (0·27–2·13) P_m = 0·60†	1/461 (0·2)	0·35 (0·04–3·42) P_m = 0·37†	8/461 (1·7)	0·69 (0·24–2·03) P_m = 0·50†	3/461 (0·7)	0·64 (0·15–2·68) P_m = 0·54†
N95 non-fit-tested	16/488 (3·3)	0·48 (0·24–0·98) P_m = 0·045†	2/488 (0·4)	0·67 (0·11–4·03) P_m = 0·66†	5/488 (1)	0·39 (0·12–1·22) P_m = 0·11†	0/488 (0)	0
All N95	37/949 (3·9)	0·62 (0·28–1·35) P_m = 0·23†	3/949 (0·3)	0·52 (0·10–2·57) P_m = 0·42	13/949 (1·4)	0·54 (0·21–1·36) P_m = 0·19†	3/949 (0·3)	0·31 (0·07–1·32) P_m = 0·113†
Medical mask	33/492 (6·7)		3/492 (0·6)		13/492 (2·6)		5/492 (1)	

*ILI definition using fever >38 – note, this is less sensitive than laboratory-confirmed infection.
**Any respiratory virus.
***Odds Ratio – Medical group as reference. A random effect logistic model accounting for clustering was used to compute odd ratios.
†P_m: *P* value adjusted for clustering of hospitals using random effect logistic regression model.[29]
CRI, Clinical respiratory illness; ILI, influenza-like illness.

Table 3. Comparison with the convenience no-mask group

Arms	CRI		ILI*		Laboratory-confirmed virus**		Influenza	
	***N* (%)**	**OR (95% CI)*** **	***N* (%)**	**OR (95% CI)*** **	***N* (%)**	**OR (95% CI)*** **	***N* (%)**	**OR (95% CI)*** **
N95 fit-tested	21/461 (4·6)	0·58 (0·18–1·89) $P = 0·37$	1/461 (0·2)	0·19 (0·02–1·78) $P = 0·14$	8/461 (1·7)	0·55 (0·22–1·35) $P = 0·19$	3/461 (0·7)	0·52 (0·13–2·09) $P = 0·36$
N95 non-fit-tested	16/488 (3·3)	**0·36 (0·14–0·94) $P = 0·038$**	2/488 (0·4)	0·33 (0·06–1·72) $P = 0·19$	5/488 (1)	**0·33 (0·12–0·89) $P = 0·03$**	0/488 (0)	0 –
All N95	37/949 (3·9)	0·46 (0·19–1·11) $P = 0·085$	3/949 (0·3)	0·26 (0·06–1·11) $P = 0·068$	13/949 (1·4)	**0·43 (0·20–0·91) $P = 0·02$**	3/949 (0·3)	0·25 (0·06–1·00) $P = 0·051$
Medical mask	33/492 (6·7)	0·74 (0·29–1·88) $P = 0·52$	3/492 (0·6)	0·49 (0·12–2·07) $P = 0·33$	13/492 (2·6)	0·84 (0·38–1·85) $P = 0·67$	5/492 (1)	0·81 (0·25–2·68) $P = 0·73$

*ILI definition using fever >38 – note, this is less sensitive than laboratory-confirmed infection.
**Any respiratory virus.
***Odds Ratio – No-mask convenience group as reference. A random effect logistic model accounting for clustering was used to compute odd ratios.
CRI, clinical respiratory illness; ILI, influenza-like illness.
Bold text signifies statistical significance.

Table 4. Multivariable random effect logistic regression model adjusting for potential confounders, for comparison of All N95 with surgical mask

Variables in the model	Odds ratio (95% CI)*			
	CRI	**ILI**	**Laboratory-confirmed virus**	**Influenza**
AllN95	**0·38 (0·17, 0·86)**	0·58 (0·10, 3·47)	**0·19 (0·05, 0·67)**	0·27 (0·06, 1·17)
Hospital Level	**0·40 (0·17, 0·96)**	1·10 (0·10, 12·60)	**0·17 (0·05, 0·65)**	*
High-risk procedures	0·92 (0·53, 1·59)	0·54 (0·10, 2·78)	0·71 (0·30, 1·68)	1·90 (0·37, 9·75)
Flu vaccine 2008	1·18 (0·59, 2·37)	**	1·07 (0·35, 3·29)	1·97 (0·23, 16·54)
Handwashing	1·00 (0·38, 2·58)	**	**	**

*Odds ratio was estimated by using random effect logistic regression model to adjust for the clustering effect.
**Model did not converge after including the variable because of multi-collinearity.
CRI, Clinical respiratory illness; ILI, influenza-like illness.
Bold text signifies statistical significance.

intention-to-treat analysis, N95 respirators but not medical masks had significantly lower rates of infection compared to no masks. However, the convenience no-mask group was not a randomized control arm and hospitals in this group were actually selected on the basis that most of their staff did not wear masks (which is not the norm in hospitals in Beijing), suggesting that conditions in those hospitals were different than those in hospitals from the masks groups. As a consequence, it is not possible to make any definitive judgement on the efficacy of masks on this basis. One possible bias would be if those hospitals had differentially higher risk of infection compared to the intervention hospitals, for example because of the occurrence of outbreaks. However, we monitored all hospitals involved in the study for outbreaks which may have increased apparent attack rates, and none were documented. Other than that, possible sources of bias that could have plausibly increased the infection rate in the control arm (namely vaccination, handwashing, hospital level and high-risk procedures) were measured. In a *post hoc* adjusted analysis, only hospital level and the N95 arm were significant against CRI and laboratory-confirmed viral infection.

Respiratory protection is a key strategy for pandemic control and key to sustaining the health care workforce. The fact that rates of all outcomes were consistently lower in the N95 group suggest that N95 respirators might offer better protection for HCWs; but a larger trial is needed to make a definitive judgment about the relative efficacy of respirators and medical masks. A recent, smaller trial found no difference between N95 and medical masks, but was

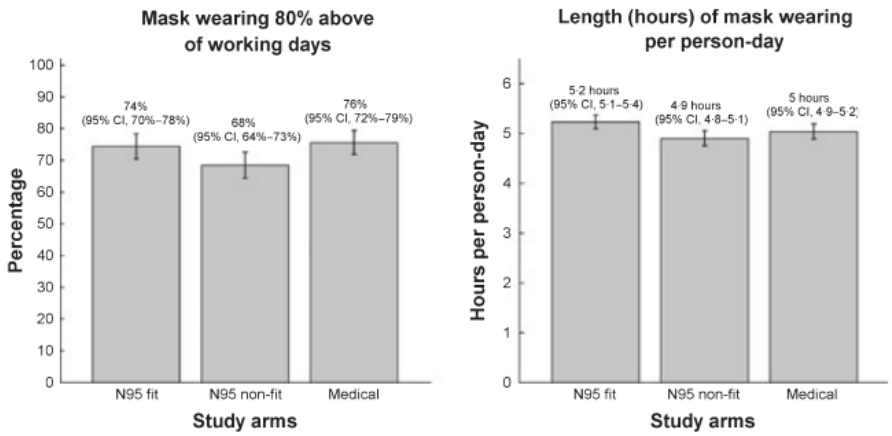

Figure 3. Rates of mask/respirator wearing (compliance defined as mask/respirator wearing ≥80% during working hours on follow-up).

Table 5. Reported problems associated with using the masks or respirators

Problems with	Medical mask (*n* = 492)	All N95 (*n* = 949)	*P* value
Using the mask/respirator			
None	85·5% (420/491)	47·4% (447/943)	<0·01
Uncomfortable	9·8% (48/491)	41·9% (395/943)	<0·01
Forgot to wear it	0% (0/491)	1·7% (16/943)	<0·01
Patient felt unconformable	0·2% (1/491)	1·8% (17/943)	0·01
Trouble communicating with the patient	3·0% (9/303)	8·0% (62/775)	<0·01
Wearing the mask/respirator			
Headaches	3·9% (11/281)	13·4% (94/701)	<0·01
Skin rash	4·6% (13/281)	5·0% (35/701)	0·81
Difficulty breathing	12·5% (35/281)	19·4% (136/701)	0·01
Allergies	9·3% (26/281)	7·1% (50/701)	0·26
Pressure on nose	11·0% (31/281)	52·2% (366/701)	<0·01
Other	0·7% (2/280)	8·3% (58/701)	<0·01

probably underpowered to detect any differences.[11] Further, the intervention in that study was use of respiratory protection only during care of identified febrile patients with ILI or high-risk procedures. This is different from the intervention in our study, which comprised wearing the mask for the entire shift. In addition, that study measured serological evidence of influenza as an outcome, which comprised the majority of outcomes, but did not exclude influenza-vaccinated participants, a flaw that would have resulted in false-positive cases of 'influenza'.

The finding that fit testing did not improve the efficacy of N95 respirators is important, although it could be explained by a lack of power. The value of fit testing varies with the quality of the respirator, and our study used a high-quality respirator. These results would not be generalizable to other respirators, where fit testing may be more important. As such, we still recommend that fit testing be part of the process of using respirators.

The small number of randomization units along with the small numbers of cases means that estimation of multivariate models would not necessarily converge. In the *post hoc* multivariable analysis, we could not adjust for all of the factors because of high correlation among some of them.

Other limitations of the study include the generalizability of our results to other types of respirators and to other HCW populations in other countries. Scoping work with Australian HCWs showed compliance of 10% with continual mask wearing during a severe influenza season.[24] Beijing was selected to maximize the power of the study because of the strong culture of mask wearing among HCWs. Another limitation of the study is that cluster RCTs can be impacted by heterogeneity of behaviours, meaning that we cannot exclude such effects caused by behaviours we did not measure. The cluster design is also strength, as interventions against infectious diseases can have herd effects. In infectious diseases which can spread from person

to person, the 'herd effect' is a real and documented phenomenon where protecting some individuals with an intervention (most commonly vaccination, but also applicable to other interventions) can also protect individuals who were not protected by the intervention. Therefore, if some individuals are randomized to masks on a ward, the individuals who do not wear masks may also be protected because of the effect the masks have on interrupting the transmission of disease from person to person. This is why it is preferable to use cluster design, where everyone in the cluster gets the same intervention.

In our study, masks or respirators were worn during the entire shift. Some policies recommend mask/respirator use only when HCWs are conducting high-risk procedures or entering an isolation room. Whether masks/respirators will be protective when used only when an identified episode of exposure occurs depends on whether HCWs accurately identify all episodes of risk, whether most transmission occurs after clearly identified exposures and whether there is transmission from asymptomatic or pre-symptomatic infections. There is currently no evidence on how much of a HCWs' risk is unidentified or unrecognized. In our study, HCWs who conducted high-risk procedures had higher rates of CRI, but not of laboratory-confirmed pathogens or influenza. Further clinical research is required to determine the efficacy of continuous versus targeted mask use.

Until now, public health policy for dealing with pandemics has relied heavily on data from a modest number of often old and inadequate studies. Data from the SARS outbreak showed that masks reduced transmission of SARS and other viral respiratory infections.[25,26] During SARS, the use of N95 respirators and medical masks was the major protective infection control measure.[27] However, the relative contribution of each type or the difference between N95 respirators and medical masks cannot clearly be determined from observational data.

Problems with adherence to mask/respirator use are also a potential problem. We showed that in Australia, less than half of parents who were randomized to wear a medical mask or respirator while their child was ill adhered with mask wearing.[10] There may be adverse effects of wearing masks, which can reduce adherence.[28–30] Our study showed significantly higher reported adverse effects of N95 respirators compared to medical masks, consistent with other studies.[28] Interestingly, this population of Chinese HCWs reported overall similar rates of discomfort with masks as parents in our household study,[10] with higher rates in the N95 group, but it did not affect their adherence with mask/respirator wearing. This suggests that discomfort is not the primary driver of adherence, and rather, cultural acceptability and other behavioural factors may be the main reason for non-adherence. The past experience of Beijing health workers with SARS may also be a factor in the high adherence. This level of adherence may not translate to Western cultural contexts in a normal winter season, especially for N95 respirators; however, adherence can change with perception of risk. During a pandemic, we would expect HCWs to have higher adherence to infection control measures. In summary, our study adds evidence on the use of respiratory protection for HCWs, but highlights the need for larger trials and comparison of different policy options.

Acknowledgements

We acknowledge the guidance and support we received from the staff at 3M China. This comprised assistance with fit testing for the study, but not financial support. Thanks also go to the trial staff at the Beijing Centre for Disease Control and Surveillance and their affiliated public health units. Thanks to the staff from the 15 Beijing hospitals which participated. Professor Raina MacIntyre designed the study, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Dr Simon Cauchemez thanks the Research Council, UK, and Neil Ferguson thanks the Medical Research Council for Centre funding.

Author contributions

Professor C. Raina MacIntyre: As the lead investigator Raina MacIntyre was responsible for conception and design of the trial, overseeing the whole study, analysing data and writing the report; Professor Quanyi Wang: Implementation, contribution to design, analysis and drafting of paper; Dr Simon Cauchemez: Statistical analysis and drafting of paper; Dr Holly Seale: Study design, form/database development, monitoring and review of paper; Professor Dominic E Dwyer: Study design, clinical and laboratory technical assistance and drafting of paper; Dr Peng Yang: Project manager; Dr Weixian Shi: Laboratory testing in China; Dr Zhanhai Gao: Statistical analysis and drafting of paper; Dr Xinghuo Pang: Recruitment and training; Dr Yi Zhang: Database management and analysis; Dr Xiaoli Wang: Database management and analysis; Dr Wei Duan: Recruitment and training; Dr Bayzidur Rahman: Statistical analysis and drafting of paper; Professor Neil Ferguson: Statistical analysis and drafting of paper.

Conflict of interests

Professor Raina MacIntyre: Raina MacIntyre receives funding from influenza vaccine manufacturers GSK and CSL Biotherapies for investigator-driven research. She has also been on advisory boards for Wyeth, GSK and Merck. Dr Simon Cauchemez received consulting fees from

Sanofi-Pasteur MSD on the modelling of varicella zoster virus. The remaining author(s) declare that they have no competing interests. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication. Prior to the start of this study, NMF acted as a consultant for Roche, Novartis and GSK Biologicals (ceasing in 2007).

References

1 Fraser C, Donnelly CA, Cauchemez S *et al.* Pandemic potential of a strain of influenza A (H1N1): early findings. Science 2009; 324(5934):1557–1561.
2 Berg HF, Van Gendt J, Rimmelzwaan GF, Peeters MF, Van Keulen P. Nosocomial influenza infection among post-influenza-vaccinated patients with severe pulmonary diseases. J Infect 2003; 46:129–132.
3 Sartor C, Zandotti C, Romain F *et al.* Disruption of services in an internal medicine unit due to a nosocomial influenza outbreak. Infect Control Hosp Epidemiol 2002; 23:615–619.
4 Horcajada JP, Pumarola T, Martinez JA *et al.* A nosocomial outbreak of influenza during a period without influenza epidemic activity. Eur Respir J 2003; 21(2):303–307.
5 Donnelly CA, Ghani AC, Leung GM *et al.* Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. Lancet 2003; 361(9371):1761–1766.
6 Larson E, Aiello A, Lee LV *et al.* Short- and long-term effects of handwashing with antimicrobial or plain soap in the community. J Community Health 2003; 28(2):139–150.
7 Larson EL, Lin SX, Gomez-Pichardo C *et al.* Effect of antibacterial home cleaning and handwashing products on infectious disease symptoms: a randomized, double-blind trial. Ann Intern Med 2004; 140(5):321–329.
8 Rabie T, Curtis V, Rabie T, Curtis V. Handwashing and risk of respiratory infections: a quantitative systematic review. Trop Med Int Health 2006; 11(3):258–267.
9 Cowling BJ, Chan KH, Fang VJ *et al.* Facemasks and hand hygiene to prevent influenza transmission in households: a randomized trial. Ann Intern Med 2009; 151:437–446.
10 MacIntyre C, Cauchemez S, Dwyer DE *et al.* Face mask use and control of respiratory virus transmission in households. Emerg Infect Dis 2009; 15:233–241.
11 Loeb M, Dafoe N, Mahony J *et al.* Surgical mask vs N95 respirator for preventing influenza among health care workers: a randomized trial. JAMA 2009; 302(17):1865–1871.
12 Shine KI, Rogers B, Goldfrank LR. Novel H1N1 Influenza and Respiratory Protection for Health Care Workers. N Engl J Med 2009; 361(19):1823–1825.
13 Balazy A, Toivola M, Adhikari A, Sivasubramani SK, Reponen T, Grinshpun SA. Do N95 respirators provide 95% protection level against airborne viruses, and how adequate are surgical masks? Am J Infect Control 2006; 34(2):51–57.
14 Lawrence RB, Duling MG, Calvert CA *et al.* Comparison of performance of three different types of respiratory protection devices. J Occup Environ Hyg 2006; 3(9):465–474.
15 Weber A, Willeke K, Marchioni R *et al.* Aerosol penetration and leakage characteristics of masks used in the health care industry. Am J Infect Control 1993; 21(4):167–173.
16 Orr NW. Is a mask necessary in the operating theatre? Ann R Coll Surg Engl 1981; 63(6):390–392.
17 Mitchell NJ, Hunt S. Surgical face masks in modern operating rooms–a costly and unnecessary ritual? J Hosp Infect 1991; 18(3):239–242.
18 Tunevall TG. Postoperative wound infections and surgical face masks: a controlled study. World J Surg 1991; 15(3):383–387.
19 Carrat F, Sahler C, Rogez S *et al.* Influenza burden of illness: estimates from a national prospective survey of household contacts in France. Arch Intern Med 2002; 162(16):1842–1848.
20 Rabe-Hesketh S, Skrondal A. Multilevel and Longitudinal Modeling Using Stata, 2nd edn. College Station, Tex: Stata Press Publication, 2008.
21 Vittinghoff E. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. New York: Springer, 2005.
22 Agresti A. An Introduction to Categorical Data Analysis, 2nd edn. Hoboken, NJ: Wiley-Interscience, 2007.
23 Donner A, Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. London: Arnold, 2000.
24 Seale H, Corbett S, Dwyer DE, Macintyre CR. Feasibility Exercise to Evaluate the Use of Particulate Respirators by Emergency Department Staff During the 2007 Influenza Season. Infect Control Hosp Epidemiol 2009; 30:710–712.
25 Lo JY, Tsang TH, Leung YH, Yeung EY, Wu T, Lim WW. Respiratory infections during SARS outbreak, Hong Kong, 2003. Emerg Infect Dis 2005; 11(11):1738–1741.
26 Pang X, Zhu Z, Xu F *et al.* Evaluation of control measures implemented in the severe acute respiratory syndrome outbreak in Beijing, 2003. JAMA 2003; 290(24):3215–3221.
27 Seto WH, Tsang D, Yung RW *et al.* Effectiveness of precautions against droplets and contact in prevention of nosocomial transmission of severe acute respiratory syndrome (SARS). Lancet 2003; 361(9368):1519–1520.
28 Lim EC, Seet RC, Lee KH *et al.* Headaches and the N95 face-mask amongst healthcare providers. Acta Neurol Scand 2006; 113(3):199–202.
29 Kao TW, Huang KC, Huang YL *et al.* The physiological impact of wearing an N95 mask during hemodialysis as a precaution against SARS in patients with end-stage renal disease. J Formos Med Assoc 2004; 103(8):624–628.
30 Troendle JF. A Stepwise Resampling Method of Multiple Hypothesis-Testing. J Am Stat Assoc 1995; 90(429):370–378.