CVI Journals.ASM.org

Meta-Analysis and Potential Role of Preexisting Heterosubtypic Cellular Immunity Based on Variations in Disease Severity Outcomes for Influenza Live Viral Challenges in Humans

Olga Pleguezuelos, Stuart Robinson, Ana Fernandez, Gregory A. Stoloff, Wilson Caparrós-Wanderley

SEEK, Central Point, London, United Kingdom

Influenza live viral challenges in humans are valuable models for testing the efficacy of vaccines and antiviral agents. Volunteers are treated with an investigational agent, and their clinical outcomes postchallenge are compared to those of placebo-treated volunteers. Despite using a common protocol, similar recruitment criteria, and similar doses of the same challenge strain, we noticed differences in disease severity outcomes between the placebo groups from different studies. We investigated whether these differences were significant and, if so, whether any pattern and its possible causes could be identified. We compared the clinical outcomes postchallenge in placebo groups from five clinical studies carried out between 2008 and 2013. Correlations between the prechallenge heterosubtypic cellular response (gamma interferon [IFN-γ]) and postchallenge clinical outcomes were also investigated in one study. Placebo groups from studies carried out between 2009 and 2010 attained significantly reduced ($P < 0.05$) symptom scores postchallenge compared to those of placebo groups from studies carried out in either 2008 or 2013. Also, in a 2010 study, the frequency of high-influenza heterosubtypic cellular responders prevaccination was significantly lower in the test group (FLU-v) than that in the placebo group ($P = 0.04$). Moreover, the increased preexisting heterosubtypic cellular response of the placebo group correlated with reductions in symptom score and viral shedding postchallenge ($P \leq 0.023$). Only postvaccination did the test group display an equivalent correlation. The last influenza pandemic coincided with a significant reduction in disease severity outcomes. This reduction also appears to correlate with increased preexisting influenza heterosubtypic cellular responses. (This study is registered at ClinicalTrials.gov under registration number NCT01226758.)

Influenza live viral challenges in humans are valuable models for testing the efficacy of vaccines and antiviral agents. Their basis is simple: a group of volunteers is treated with an investigational agent, and their clinical outcomes postchallenge are compared to those of a group of placebo-treated volunteers. Their logistics, in contrast, are complex.

Influenza infection elicits a range of immune responses. One such response is the production of strain-specific neutralizing antibodies that confer immunity against infection by the same strain (1). As a result, a key volunteer exclusion criterion in challenge studies is the detection of preexisting neutralizing antibodies (hemagglutination inhibition [HAI], >10) to the challenge strain. Another such response is the generation of antiviral cellular immune responses. Despite existing evidence as to their protective role during infection (2–4), preexisting cellular immune responses to the challenge strain are not normally assessed during volunteer recruitment.

We have developed a novel vaccine (FLU-v) that elicits broad influenza heterosubtypic cellular responses without inducing any significant antibody response (5–7). In humans, FLU-v was found to be safe and well tolerated and, in a live viral challenge study, to induce a vaccine-specific cellular response whose magnitude correlated with reductions in symptom score and viral shedding (7). No such correlations were seen in the placebo group, but we did notice that both viral shedding and symptom score postchallenge were much lower (50%) in our placebo group than those in the placebo group from a previous study. To establish the significance of these differences, we compared the placebo group outcomes of several other influenza live viral challenge studies. All these studies, although involving different placebo agents, were carried out by the same clinical group (Retroscreen Ltd.), using the same recruitment criteria, viral strain and dose, and method for determining postchallenge clinical and virological outcomes. This meta-analysis revealed an "experiment of nature" that we believe provides interesting insights in the potential of the cellular immune system for controlling influenza virus infection.

MATERIALS AND METHODS

Clinical trial data used for meta-analysis. The reported postchallenge clinical outcomes for the placebo group of four reported independent clinical trials (3, 7–9) and one previously unreported study (Retroscreen Ltd., personal communication) were used for the meta-analysis. The placebo agents used in the studies were different, but all the studies were carried out by the same clinical group (Retroscreen Ltd.) and were conducted according to a common challenge protocol (Fig. 1) that used the same well-defined recruitment criteria, viral challenge strain (A/Wisconsin/67/2005, H3N2), and procedures for the assessment of disease

Received 23 February 2015 **Returned for modification** 1 April 2015
Accepted 14 June 2015

Accepted manuscript posted online 17 June 2015

Citation Pleguezuelos O, Robinson S, Fernandez A, Stoloff GA, Caparrós-Wanderley W. 2015. Meta-analysis and potential role of preexisting heterosubtypic cellular immunity based on variations in disease severity outcomes for influenza live viral challenges in humans. Clin Vaccine Immunol 22:949–956. doi:10.1128/CVI.00101-15.

Editor: D. W. Pascual

Address correspondence to Wilson Caparrós-Wanderley, wilson.wanderley@seekacure.com.

doi:10.1128/CVI.00101-15

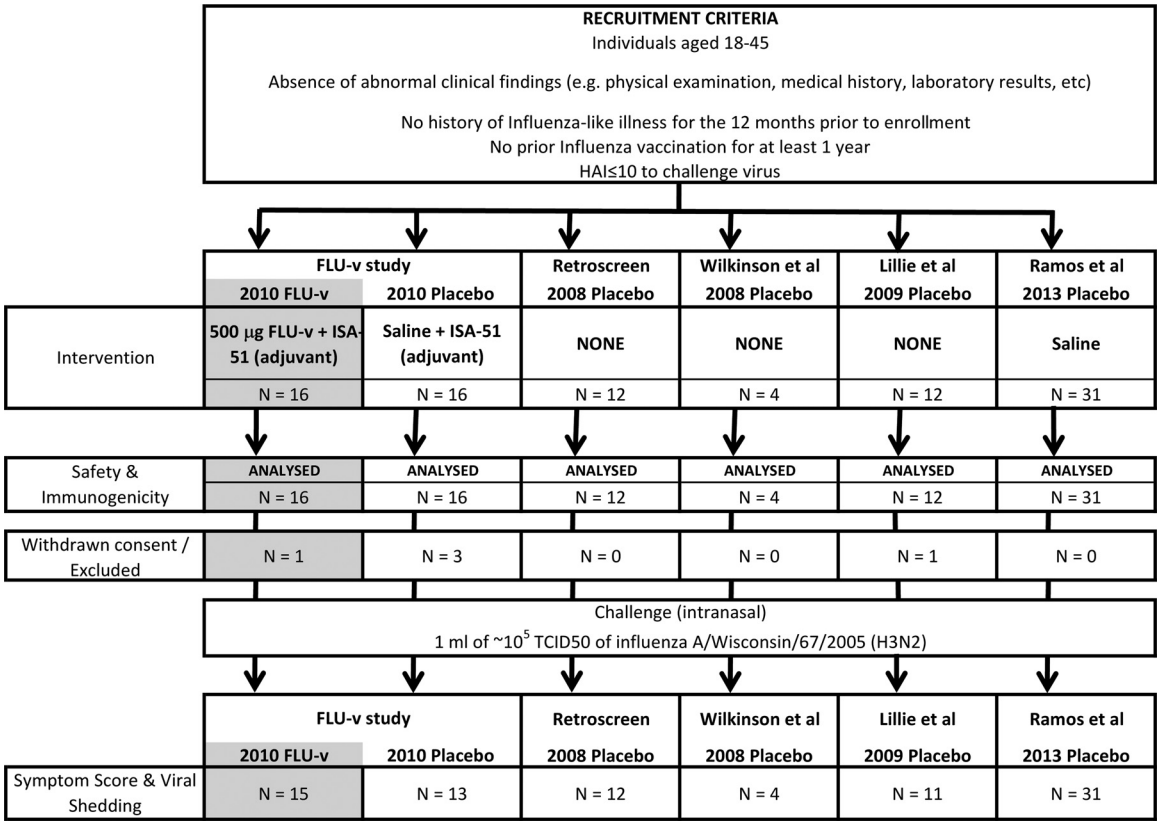

FIG 1 Consort profile. Shown are the trial profile and baseline demographic data for enrolled volunteers in all five studies analyzed. The reported median age of the volunteers in the studies ranged from 24 to 30 years. Where this information is provided, studies are reported to have been carried out between August and November. The section in gray refers to data not incorporated in the meta-analysis of the placebo groups but used for the comparison of cellular immunity described later in the paper.

severity and viral shedding. The exact details for each study are provided in the reports listed above, but they are also briefly summarized below.

Recruitment criteria and study procedures. Healthy male subjects age 18 to ~45 years with no clinically significant abnormal findings (i.e., physical examination, medical history, or laboratory results) and no medical history of influenza-like illness in the prior 12 months were assessed for enrollment. Only those with an HAI of ≤10 for the influenza challenge strain were enrolled.

Following recruitment and treatment (placebo or test agent), volunteers were challenged on day 0 by nasal instillation with 1 ml of solution containing approximately $10^{5.25}$ 50% tissue infective dose per ml of live A/Wisconsin/67/2005 (H3N2) (tissue culture grown). From days 5 to 7, the volunteers received antiviral treatment (e.g., oseltamivir) before being released from quarantine on day 7.

Physical examinations and clinical laboratory tests were performed at screening, pre- and posttreatment (both prechallenge), and daily from day −2 prechallenge to day 7 postchallenge. A final assessment was carried out around day 28 postchallenge. Volunteer self-recorded observations pre- and postchallenge and the scripted symptom questionnaires were assessed by clinical staff.

Symptom scoring and virology and HAI tests. The symptom score was determined using a standardized scoring system (3, 10) based on subject self-assessment and examination by a clinician. A range of parameters (e.g., runny or stuffy nose, sneezing, sore throat, earache, malaise, cough, shortness of breath, headache, and muscle/joint ache) were assessed and scored from 0 (absent) to 3 (severe).

Viral shedding in the nasopharyngeal samples was determined by a 50% tissue culture infective dose ($TCID_{50}$) assay, as described in the WHO Manual for the Laboratory Diagnosis and Virological Surveillance of Influenza (11). Briefly, serial 10-fold dilutions of virus-containing daily postchallenge nasal lavage samples were inoculated into 96-well microtiter plates seeded with Madin-Darby canine kidney (MDCK) cells. Cytopathic effects in individual wells were determined after 5 to 6 days of incubation at 37°C. Viral shedding was defined as a viral culture titer of >1.5 $\log_{10}$ $TCID_{50}$/ml.

Hemagglutinin (HA)-specific antibody titers against the challenge virus in volunteer serum samples were determined by HAI assay using chicken erythrocytes, as described in the WHO Manual for the Laboratory Diagnosis and Virological Surveillance of Influenza (11).

Regulatory approval and ethical considerations. All studies included in the meta-analysis were reported as conducted in accordance with good clinical practice, the Declaration of Helsinki (1964 and 2008), and all regulatory requirements.

As we are also reporting previously undisclosed experimental data, we confirm that our FLU-v study (7) was approved by the Plymouth Independent Ethics Committee under REC reference number 10/IEC04/1. The trial was registered under European Clinical Trials database (EudraCT) identifier 2009-014716-35 and registered at ClinicalTrials.gov under registration no. NCT01226758. Written informed consent was obtained from all participants.

Vaccine description. FLU-v is a sterile equimolar mixture of four polypeptides encoding immunoreactive conserved regions within the influenza virus (5–7). These sequences were synthetically manufactured (Bachem AG, Bubendorf, Switzerland) in accordance with current good manufacturing practice: M1, DLEALMEWLKTRPILSPLTKGILGFVFT LTVP (32 amino acids [aa]); NPA, DLIFLARSALILRGSVAHKSC (21 aa);

NPB, PGIADIEDLTLLARSMVVVR (20 aa); and M2, IIGILHLILWILDR LFFKCIYRLF (24 aa).

FLU-v was administered subcutaneously in a 1.0-ml volume as a single 500-μg dose in saline emulsified (1:1) with adjuvant ISA-51 (Seppic, France). The placebo was saline emulsified with ISA-51. The adjuvant is composed of a light mineral oil and a surfactant system designed to make a water-in-oil emulsion. Functionally, ISA-51 is not known to preferentially favor the induction of Th1-like responses (12).

Heterosubtypic cellular immunity-cytokine ELISA. Blood was harvested prechallenge on days −21 (i.e., prevaccination) and −2 (i.e., 19 days postvaccination), and peripheral blood mononuclear cells (PBMCs) were isolated and frozen. Thawed PBMCs were seeded at 2×10^5 cells/well (96-well plate) in RPMI 1640 (Sigma, United Kingdom), supplemented with 25 mM HEPES, penicillin (100 units/ml), streptomycin (100 μg/ml), 10% fetal calf serum (FCS), and one of the following test antigens: 1 μg/ml concanavalin A (ConA) (Sigma), 1 μg/ml bovine serum albumin (BSA) (Sigma), or live influenza A/Swine/Iowa/15/30 (H1N1) (multiplicity of infection, 10). Virus (egg grown) was obtained from the National Institute for Biological Standards and Control (NIBSC) as low-endotoxin preparations suitable for *in vitro* cellular analysis. Each antigen was tested in triplicate. After 24 h of incubation at 37°C and 5% CO_2, gamma interferon (IFN-γ) production in the cell supernatant for each of the test antigens was determined using a validated enzyme-linked immunosorbent assay (ELISA) (human IFN-γ kit 555142; BD, United Kingdom). The response levels were calculated as picograms per milliliter of IFN-γ produced against a standard provided with the assay kit. The minimum level of detection for the assay is 9 pg/ml IFN-γ.

Strong heterosubtypic cellular responses were defined as those in which an IFN-γ response of an individual to the influenza A/Swine/Iowa/15/30 (H1N1) virus was ≥4-fold higher than the IFN-γ response of the individual to the negative control (i.e., BSA plus medium).

Statistical analysis. Intergroup differences in total mean symptom score and viral shedding were determined by single-factor analysis of variance (ANOVA). Pairwise differences between the studies were determined by *t* test (2-way), with an adjustment of significance for multiple pairwise comparisons made using the Tukey-Kramer honestly significant difference (HSD) method. Heterosubtypic responder frequencies were analyzed by the Friedman exact test, while correlations between clinical outcomes and heterosubtypic cellular response levels were determined using the Spearman rank correlation test.

RESULTS

Mean total symptom score postchallenge and interstudy variability. We previously reported (7) how in an influenza live viral challenge study carried out in 2010, vaccination with FLU-v induced an IFN-γ response to the vaccine, the magnitude of which correlated with reductions in both viral titer ($P = 0.01$) and total symptom score ($P = 0.02$). No such correlation was seen in the placebo group. Although we saw no significant differences in mean total symptom score postchallenge between the FLU-v and placebo group, we did notice a significant reduction in mean total symptom score postchallenge in our 2010 placebo group compared to that of the placebo group of a previous unreported study carried out by Retroscreen in 2008 (Retroscreen Ltd., personal communication) (mean ± standard deviation [SD] total symptom score, 11.4 ± 13.0 versus 37.1 ± 27.5 for our placebo versus 2008 placebo; $P = 0.006$).

This significant difference in outcomes was surprising to us, because, although the natures of the placebo agent were different in the two trials, the historical 2008 placebo data set ($n = 12$) had been obtained by the same clinical group (Retroscreen Ltd.), using the same recruitment criteria, viral strain and dose, and method for determining the symptom score. More importantly, the outcomes of this 2008 placebo group constituted the baseline data used to calculate the sample size required to meet the endpoints of our trial.

This difference in outcomes also raised the question of whether our observation was unique or whether significant differences in outcome were a common observation in live viral challenge studies. To address this question, we compared the outcome of both our 2010 placebo group and the 2008 Retroscreen (Retroscreen Ltd., personal communication) placebo group against those reported for placebo groups in other published studies carried out in 2008 (3), 2009 (9), and 2013 (8). These studies were performed according to the same common standard protocol (Fig. 1), recruitment criteria, and procedures used in the 2010 and 2008 placebo groups.

Statistical analysis (ANOVA) of these five studies (Table 1) revealed a significant difference ($P = 0.004$) in the mean total symptom score of the placebo groups. Subsequent pairwise comparisons (*t* test with Tukey-Kramer's HSD adjustment for significance) revealed that the mean total symptom score for the placebo group in the 2008 Wilkinson et al. study (3) was significantly higher than that seen in our 2010 placebo group (mean ± SD, 60.8 ± 10.7 versus 11.4 ± 13.0, respectively; $P = 0.000$), but not different from that in the 2008 Retroscreen (Retroscreen Ltd., personal communication) placebo group (mean ± SD, 60.8 ± 10.7 versus 37.1 ± 27.5, respectively; $P > 0.050$). In contrast, the mean total symptom score for the placebo group in the 2009 Lillie et al. study (9) was significantly lower than that seen in both the 2008 Retroscreen (Retroscreen Ltd., personal communication) placebo group (mean ± SD, 15.3 ± 15.1 versus 37.1 ± 27.5, respectively; $P = 0.030$) and the 2008 Wilkinson et al. (3) placebo (mean ± SD, 15.3 ± 15.1 versus 60.8 ± 10.7, respectively; $P = 0.000$) but not different from that in our 2010 placebo group (mean ± SD, 15.3 ± 15.1 versus 11.4 ± 13.0, respectively; $P > 0.050$). A final comparison of these four different placebo groups with the placebo group in the 2013 Ramos et al. study (8) (mean ± SD, 55.5 ± 54.8) reveals that mean total symptom score in this study is not different from that seen in either the 2008 Wilkinson et al. placebo group (3) or the 2008 Retroscreen (Retroscreen Ltd., personal communication) placebo group ($P > 0.050$ for both), but it is significantly higher than that seen in both the 2009 Lillie et al. (9) placebo group ($P = 0.022$) and our 2010 placebo group ($P = 0.007$). These results indicate that following influenza live viral challenge, the mean total symptom scores in placebo group volunteers were significantly lower in 2009 to 2010 than they were in either 2008 or 2013.

Mean total viral shedding postchallenge and interstudy variability. We then proceeded to test whether the observed differences in mean total symptom score across the studies were also reflected in the mean total viral shedding measurements. Total viral shedding data were not reported in the Wilkinson et al. study (3), and hence, we did not include this study in the analysis. Nonetheless, a comparison of the remaining four studies revealed a significant difference ($P = 0.040$) in mean total viral shedding.

Subsequent pairwise analysis revealed that, as shown in Table 1, mean total viral shedding in the 2008 Retroscreen (Retroscreen Ltd., personal communication) placebo group (mean ± SD, 10.1 ± 2.9) was significantly higher than in the 2009 Lillie et al. (9) placebo group (mean ± SD, 3.3 ± 4.3; $P = 0.012$), our 2010 placebo group (mean ± SD, 4.0 ± 4.4; $P = 0.022$) and the Ramos 2013 (8) placebo group (mean ± SD, 3.2 ± 4.5; $P = 0.006$).

TABLE 1 Summary of descriptive statistics for the postchallenge outcomes in all analyzed studies[d]

Data by outcome	Result for placebo group in[a]: Retroscreen 2008 ($n = 12$)	Wilkinson et al. 2008 ($n = 4$)	Lillie et al. 2009 ($n = 11$)	FLU-v study 2010 ($n = 13$)	Ramos et al. 2013 ($n = 31$)
Total symptom score[b]					
Avg	37.1	60.8	15.3	11.4	55.5
SD	27.5	10.7	15.1	13.0	54.8
Median	50.0	57.5	8.0	8.0	39.0
Minimum	0.0	52.0	0.0	0.0	0.0
Maximum	74.0	76.0	38.0	44.0	190.0
Pairwise comparison *t* test (*P* value)					
Retroscreen 2008					
Wilkinson et al. 2008	>0.050				
Lillie et al. 2009	0.030	0.000			
FLU-v study 2010	0.006	0.000	>0.050		
Ramos et al. 2013	>0.050	>0.050	0.022	0.007	
Infection rate (%)[c]	66.7	100.0	45.5	61.5	48.4
Total viral shedding					
Avg	10.1	NA[e]	3.3	4.0	3.2
SD	2.9	NA	4.3	4.4	4.5
Median	10.6	NA	0.0	2.8	0.0
Minimum	6.5	NA	0.0	0.0	0.0
Maximum	12.5	NA	10.8	12.5	14.3
Pairwise comparison *t* test (viral shedding) (*P* value)					
Retroscreen 2008					
Wilkinson et al. 2008	NA				
Lillie et al. 2009	0.012	NA			
FLU-v study 2010	0.022	NA	>0.050		
Ramos et al. 2013	0.006	NA	>0.050	>0.05	

[a] The references for the studies are Retroscreen 2008, personal communication; Wilkinson et al. 2008, 3; Lillie et al. 2009, 9; FLU-v 2010, 7; and Ramos et al. 2013, 8.
[b] Total symptom score is the sum of all measured symptoms scores for an individual from day 1 to day 7 following challenge with influenza A/Wisconsin/67/2005 (H3N2). The ANOVA *P* value was 0.004.
[c] The infection rates are the percentage of challenged volunteers with at least one daily nasal sample positive for influenza A/Wisconsin/67/2005 (H3N2) postchallenge.
[d] Total viral shedding represents the sum of all measured viral shedding for an individual from day 1 to day 5 postchallenge with influenza A/Wisconsin/67/2005 (H3N2). Viral shedding on days 6 and 7 postchallenge was not considered, as under the clinical protocol used, all individuals receive antiviral treatment (e.g., oseltamivir) on those days. The ANOVA *P* value was 0.040.
[e] NA, no data are available.

Heterosubtypic immunity. In an attempt to determine the possible reasons for the differences among the groups, we first analyzed the infection rate for each of the studies. Infection rate was defined as the percentage of volunteers with at least one positive result by $TCID_{50}$ between days 1 and 5 after influenza live viral challenge. As shown in Table 1, and despite the wide range of values, no statistical differences ($P > 0.05$) were found in the infection rates across the different studies: the 2008 Wilkinson et al. (3) placebo group (100%), the 2008 Retroscreen (Retroscreen Ltd., personal communication) placebo group (66.6%), the 2009 Lillie et al. (9) placebo group (45.5%), our 2010 placebo group (61.5%), and the 2013 Ramos (8) placebo group (48.4%).

The similarity in infection rates across the studies suggests that the mechanism responsible for the differences in outcomes is most likely a postinfection mechanism. If correct, this would exclude neutralizing-antibody responses but not cellular immune responses. Unfortunately, cellular responses to the challenge virus were not assessed in any of these studies, either pre- or postchallenge. Moreover, if any cellular responses were measured, the antigen (e.g., virus or vaccine) and the method of analysis (e.g., ELISA or enzyme-linked immunosorbent spot assay [ELISPOT]) used were all different, thus rendering any direct comparison impossible.

As stated earlier, we previously established (7) that cellular responses to a vaccine correlated with reductions in both viral load and symptom score. Since the period of reduced total mean symptom scores and viral shedding (2009 to 2010) identified from our earlier meta-analysis coincided with the dates of the last influenza pandemic, we decided to test if in our study, strong preexisting heterosubtypic cellular responses to a H1N1 swine influenza strain were common and, if so, whether their intensity negatively correlated with symptom score and viral shedding. Ideally, we would have preferred to use the pandemic influenza A/California/7/2009 (H1N1) strain, but the WHO recommends the use of biosafety level 2 plus (BSL-2 plus) facilities with biosafety level 3 (BSL-3) practices with this strain (13). As these facilities are not available to us, we settled for a BSL-2 swine strain, A/Swine/Iowa/15/30 (H1N1).

TABLE 2 Heterosubtypic cellular immune responses pre- and postvaccination in 2010 FLU-v study

Value by vaccination time	Fold increase in A/Swine/Iowa/15/30 IFN-γ response compared to negative control for group[a]:			
	Placebo ($n = 7$)		FLU-v ($n = 8$)	
	Prevaccination	Postvaccination	Prevaccination	Postvaccination
Median	7.0	10.3	3.4	5.2
Minimum	1.3	3.2	1.3	1.1
Maximum	13.0	12.1	31.8	33.5

[a] Values are represented as the fold increase in IFN-γ response to A/Swine/Iowa/15/30 (H1N1) compared to the negative control. The mean ± SD IFN-γ (pg/ml) responses to the negative control pre- and postvaccination for both groups are 99.2 ± 25.7 versus 80.6 ± 12.9 pg/ml, respectively. The IFN-γ (pg/ml) response to the positive control (ConA) pre- and postvaccination for both groups are 311 ± 85 versus 378 ± 35 pg/ml, respectively.

In our 2010 study, we dosed and challenged a total of 28 volunteers. However, for this *post hoc* analysis, frozen PBMC samples were available from only 15 volunteers (seven from the placebo group and eight from the FLU-v group). We found (Table 2) strong prevaccination IFN-γ responses to the recall A/Swine/Iowa/15/30 (H1N1) virus (i.e., ≥4-fold increase in the IFN-γ response to negative control) in all but one of the placebo-treated volunteers (median, 7.0-fold increase). In contrast, in the vaccinated (FLU-v) group, a strong prevaccination IFN-γ response to the recall A/Swine/Iowa/15/30 (H1N1) virus (median, 3.4-fold increase) was found in only one volunteer. Postvaccination, the frequencies of strong IFN-γ responders became similar in the two groups (5 versus 4 in the placebo versus FLU-v group, respectively), but the overall level of IFN-γ response to the recall swine virus remained higher in the placebo group (median, 10.3 versus 5.2 in the placebo versus FLU-v group, respectively).

A correlation analysis revealed a significant negative correlation in the placebo group between the intensity of the heterosubtypic IFN-γ response to the A/Swine/Iowa/15/30 (H1N1) virus and both the mean total symptom score ($r = -0.771$, $P = 0.036$; Fig. 2A) and the mean total viral shedding ($r = -0.768$, $P = 0.022$; Fig. 2B). In the FLU-v group, no significant correlations were seen prevaccination ($P > 0.05$), but a significant negative correlation was established postvaccination between the intensity of the heterosubtypic IFN-γ response to the A/Swine/Iowa/15/30 (H1N1) virus and the mean total symptom score ($r = -0.667$, $P = 0.035$; Fig. 2C).

DISCUSSION

Influenza infection elicits a range of natural antibody and cellular immune responses to the virus. Some of these responses are specific to the infecting viral strain (homosubtypic responses), while others are cross-reactive to other viral strains (heterosubtypic responses). Among the homosubtypic responses, neutralizing antibodies directed to the hemagglutinin (HA) and neuraminidase (NA) antigens are of particular interest. Infection by one influenza strain elicits neutralizing HA/NA antibody responses that confer immunity against infection by the same strain (1). For >50 years, influenza public health programs worldwide have built upon this observation by using vaccines that induce homosubtypic HA/NA neutralizing-antibody responses. Heterosubtypic responses, despite increasing evidence of their potential protective role during infection at both the antibody level (14–16) and the cellular level (2–4), have not yet been successfully exploited in the clinic.

Notwithstanding their universal use, HA/NA-based vaccines suffer from major shortcomings. As new variants of the virus emerge every year, the new circulating viral strains must be first identified before new formulations of the vaccine are prepared every year, which in turn means that every year, the population must be revaccinated (17). A clear need remains for a vaccine that can address these shortcomings.

Live viral challenge studies in humans are valuable models for the development of effective therapies (e.g., vaccines and antivirals) against influenza virus. They allow the efficacy of a candidate treatment to be assessed by comparing disease severity outcomes postchallenge between volunteer groups treated with either the candidate therapy or a placebo.

Recognizing the importance of neutralizing-antibody responses in influenza protection, the identification of preexisting neutralizing-antibody titers (i.e., HAI, >10) to the challenge strain is a key universal exclusion criterion during volunteer recruitment in live viral challenge studies (3, 7–9). In contrast, neither preexisting heterosubtypic immune responses (antibody or cellular) nor preexisting homosubtypic cellular responses to the challenge strain are regularly assessed during volunteer recruitment.

In 2010, we carried out a live viral challenge study in humans using a novel vaccine (FLU-v) designed to elicit cellular immune responses against influenza virus. Despite induction of a FLU-v-specific IFN-γ response (6, 7) that correlated with reductions in viral shedding and symptom score (7), no significant differences in clinical outcome were seen between the placebo and FLU-v groups. However, we did identify a clear and significant reduction in both viral shedding and symptom score in our 2010 placebo group compared to a placebo group from a study carried out by Retroscreen in 2008 (Retroscreen Ltd., personal communication). This 2008 placebo group is significant because its outcomes constituted the baseline data used to calculate the sample size needed to meet the endpoints of our 2010 trial.

As a certain degree of variability is expected in all biological systems, we decided to investigate how consistent viral shedding and symptom score outcomes were across five live viral challenge studies carried out between 2008 and 2013. Meta-analysis of historical data is extensively used in clinical research (18–23) and, under certain stringent rules, is even allowed by both the FDA and the European Medicines Agency (EMEA) to assess the efficacy of a treatment (24, 25). These rules state that all the data analyzed must come from clinical trials that used the same eligibility criteria, measured comparable variables, and were carried out by the same clinical investigators. Since all five clinical studies considered in our meta-analysis were carried out by the same clinical group (Retroscreen Ltd.), using the same well-defined recruitment criteria, were performed according to a common challenge protocol that used similar doses of the same viral strain (A/Wisconsin/67/

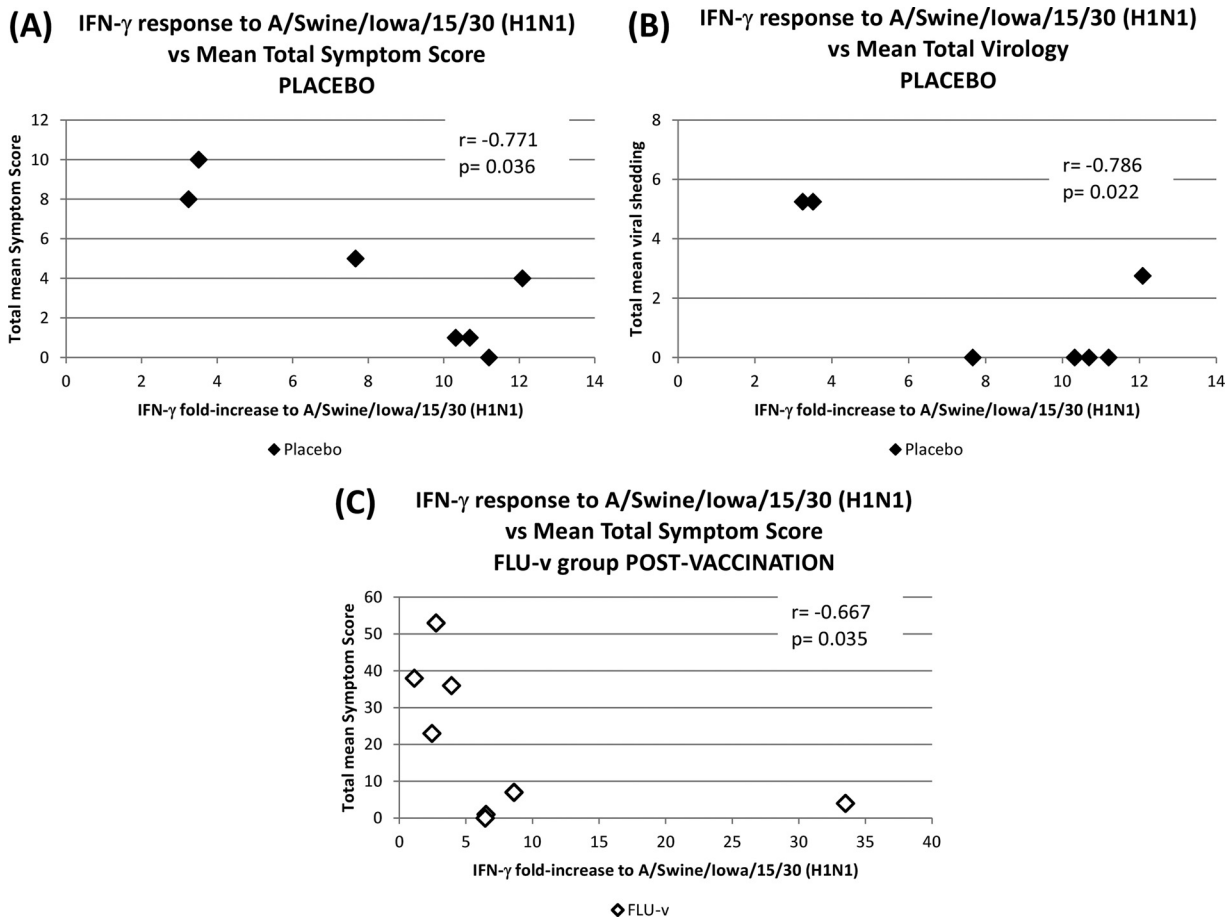

FIG 2 Correlation analysis between heterosubtypic cellular responses and measurements of disease severity postchallenge. The values are presented as the fold increase in IFN-γ response to A/Swine/Iowa/15/30 (H1N1) compared to the negative control. (A and B) Correlations between the preexisting heterosubtypic cellular response of the placebo group and its mean total symptom score and mean total viral shedding postchallenge. (C) Correlation between the postvaccination heterosubtypic cellular response of the FLU-v group and its mean total symptom score postchallenge. All analyses were carried out using the Spearman rank correlation test.

2005 [H3N2]), and assessed the same parameters (i.e., symptom score and viral shedding), we were confident of the validity of our approach. Of course, the natures of the placebo in these five studies were different, but since we and Retroscreen Ltd. agreed to use historical data from the 2008 placebo group to determine the required sample size of our 2010 study, we believe this decision was consistent with and supports our multistudy comparative approach.

The meta-analysis revealed that of the five studies analyzed, the two studies carried out between 2009 and 2010 (7, 9) achieved total mean symptom scores postchallenge that were significantly lower (~50%) than those seen in studies carried out in either 2008 (3) or 2013 (8). Viral shedding was also significantly higher in the 2008 studies than in the 2009 and 2010 studies, and, in contrast to symptom score, it was also higher than that in the 2013 study.

An accurate determination of the mechanism(s) responsible for these differences was not possible, as the immune/pharmacological effector mechanisms assessed were different for each study. However, because (i) infection rates (determined as the percentage of challenged volunteers that develop a positive $TCID_{50}$ between days 1 and 5 postchallenge) across all five studies were not statistically different, (ii) neutralizing antibodies act primarily at the preinfection stage, and (iii) all volunteers had HAI titers to the challenge strain of ≤10, we do not believe that the observed interstudy differences were caused by differences in the HAI titers of the volunteers.

An assessment of cellular responses among the studies was not possible. Preexisting cellular responses to the challenge virus are not regularly assessed in these studies and, when cellular responses are measured, the antigens (e.g., virus or vaccine) and the methods of analysis (e.g., ELISA or ELISPOT) used are different, thus rendering any direct comparison impossible. Nonetheless, three observations lead us to consider the possibility that differences in the preexisting influenza heterosubtypic cellular responses may be at least partially responsible for the observed interstudy differences in placebo group outcomes postchallenge. First, the two studies showing significant reductions in mean total symptom scores were those carried in 2009 and 2010. These dates coincide with the last influenza pandemic. Second, we (7) and others (3, 4) have shown significant negative correlations between the intensity of the cellular response and measurements of influenza disease severity. Third, the reduction in viral shedding, but not in either symptom score or rate of infection, in the 2013 study compared to that in the 2008 placebo group suggests that a postinfection mechanism controls viral replication.

Although a lack of data prevented us from comparing the role of heterosubtypic cellular responses across the five studies considered, we believed that some relevant evidence could still be obtained through additional testing of PBMC samples from our 2010 study. Unfortunately, we did not have a complete sample set for this *post hoc* analysis and hence, we accept that the small size of the sample (15 individuals) further limits the power of this analysis. Nonetheless, we found significant correlations between the preexisting IFN-γ responses to influenza A/Swine/Iowa/15/30 (H1N1) and reductions in both total mean symptom score and total mean viral shedding in the placebo group.

An additional and surprising finding of our analysis was that the frequency of preexisting high-IFN-γ responders to A/Swine/Iowa/15/30 (H1N1) was much higher in the placebo group than that in the vaccine (FLU-v) group. Moreover, although no significant correlation between the IFN-γ response to the influenza A/Swine/Iowa/15/30 (H1N1) strain and reduction in viral shedding was seen prevaccination in the FLU-v group, this correlation became evident postvaccination. Of course, we have no evidence that the pattern of heterosubtypic cellular responses (i.e., to A/Swine/Iowa/15/30 [H1N1]) is the same as that of the homosubtypic cellular responses (i.e., to the challenge strain A/Wisconsin/67/2005 [H3N2]). However, we maintain that it is not unreasonable to expect it to be so.

We have no explanation as to how, despite the randomization and double-blind nature of the study, our placebo group ended up with a higher number of volunteers with strong heterosubtypic cellular responses than our FLU-v group. The recruitment criteria and randomization in our study were not different from those of the other studies included in our meta-analysis. A *post hoc* analysis of preexisting HAI responses in our volunteers to the actual 2009-2010 pandemic strain (A/California/7/2009 [H1N1]) did not reveal any positive individual in either the placebo or the FLU-v group (data not shown). Nonetheless, we cannot completely rule out a difference in the exposure rate to the virus between the two groups. A report by Presanis et al. (26) suggests that the rate of asymptomatic infection in England during the pandemic (June 2009 to February 2010) was as high as 65%. With the benefit of hindsight, and since our study took place shortly after the end of the pandemic, it is our opinion that the list of exclusion criteria used (i.e., history of influenza-like illness over the previous 12 months and HAI of >10) was ill suited to prevent the recruitment of asymptomatically infected individuals.

Increased levels of influenza heterosubtypic cellular responses in the population after the pandemic might also help explain the particular results of the placebo group in the 2013 Ramos (8) study. McMichael et al. (27) showed that T-cell responses to influenza virus are detectable years after initial natural exposure, although their number declines rapidly with time. As T-cell responses are widely acknowledged to play a key antiviral role, it is possible that exposure to the challenge virus may have caused the expansion of a small pool of memory influenza virus heterosubtypic T-cell clones. The expansion of this small population may not have been sufficient to significantly reduce symptom severity (total symptom score), but it may have been able to have a negative effect on the rate of viral proliferation (total viral shedding).

In summary, we believe our results provide evidence of an unplanned "experiment of nature" that adds to the existing body of evidence on the ability of heterosubtypic cellular immunity to reduce influenza disease severity in humans (2–4). As such, it supports our efforts, and those of other groups, in developing vaccines that elicit heterosubtypic cellular immune responses against influenza virus. As to whether it constitutes sufficient evidence to justify the consistent screening of volunteers for preexisting cellular immunity to the challenge strain during recruitment, we leave that decision to any researcher planning to use influenza live viral challenge models in humans in the future.

REFERENCES

1. **Carrat F, Flahault A.** 2007. Influenza vaccine: the challenge of antigenic drift. Vaccine **25:**6852–6862.
2. **McMichael AJ, Gotch FM, Noble GR, Beare PA.** 1983. Cytotoxic T-cell immunity to influenza. N Engl J Med **309:**13–17. http://dx.doi.org/10.1056/NEJM198307073090103.
3. **Wilkinson TM, Li CK, Chui CS, Huang AK, Perkins M, Liebner JC, Lambkin-Williams R, Gilbert A, Oxford J, Nicholas B, Staples KJ, Dong T, Douek DC, McMichael AJ, Xu XN.** 2012. Preexisting influenza-specific $CD4^+$ T cells correlate with disease protection against influenza challenge in humans. Nat Med **18:**274–280. http://dx.doi.org/10.1038/nm.2612.
4. **Sridhar S, Begom S, Bermingham A, Hoschler K, Adamson W, Carman W, Bean T, Barclay W, Deeks JJ, Lalvani A.** 2013. Cellular immune correlates of protection against symptomatic pandemic influenza. Nat Med **19:**1305–1312. http://dx.doi.org/10.1038/nm.3350.
5. **Stoloff GA, Caparrós-Wanderley W.** 2007. Synthetic multi-epitope peptides identified *in silico* induce protective immunity against multiple influenza serotypes. Eur J Immunol **37:**2441–2449. http://dx.doi.org/10.1002/eji.200737254.
6. **Pleguezuelos O, Robinson S, Stoloff GA, Caparrós-Wanderley W.** 2012. Synthetic influenza vaccine (FLU-v) stimulates cell mediated immunity in a double-blind, randomised, placebo-controlled phase I trial. Vaccine **30:**4655–4660. http://dx.doi.org/10.1016/j.vaccine.2012.04.089.
7. **Pleguezuelos O, Robinson S, Fernandez A, Stoloff GA, Mann A, Gilbert A, Balaratnam G, Wilkinson T, Lambkin-Williams R, Oxford J, Caparrós-Wanderley W.** A synthetic influenza vaccine induces a cellular immune response which correlates with reduction in symptomatology and virus shedding in a randomised phase Ib live viral challenge in man. Clin Vaccine Immunol, in press.
8. **Ramos EL, Mitcham JL, Koller TD, Bonavia A, Usner DW, Balaratnam G, Fredlund P, Swiderek KM.** 2014. Efficacy and safety of treatment with an anti-M2e monoclonal antibody in experimental human influenza. J Infect Dis **211:**1038–1044.
9. **Lillie PJ, Berthoud TK, Powell TJ, Lambe T, Mullarkey C, Spencer AJ, Hamill M, Peng Y, Blais ME, Duncan CJ, Sheehy SH, Havelock T, Faust SN, Williams RL, Gilbert A, Oxford J, Dong T, Hill AV, Gilbert SC.** 2012. Preliminary assessment of the efficacy of a T-cell-based influenza vaccine, MVA-NP+M1, in humans. Clin Infect Dis **55:**19–25. http://dx.doi.org/10.1093/cid/cis327.
10. **Hayden FG, Fritz R, Lobo MC, Alvord W, Strober W, Straus SE.** 1998. Local and systemic cytokine responses during experimental human influenza A virus infection. Relation to symptom formation and host defense. J Clin Invest **101:**643–649.
11. **World Health Organization Global Surveillance Network.** 2011. Manual for the laboratory diagnosis and virological surveillance of influenza. World Health Organization, Geneva, Switzerland. http://whqlibdoc.who.int/publications/2011/9789241548090_eng.pdf.
12. **Shibaki A, Katz SI.** 2002. Induction of skewed Th1/Th2 T-cell differentiation via subcutaneous immunization with Freund's adjuvant. Exp Dermatol **11:**126–134. http://dx.doi.org/10.1034/j.1600-0625.2002.110204.x.
13. **World Health Organization.** 2009. Availability of a candidate reassortant vaccine virus for the novel influenza A (H1N1) vaccine development. World Health Organization, Geneva, Switzerland. www.who.int/csr/resources/publications/swineflu/ivr153_20090608_en.pdf.
14. **Corti D, Suguitan AL, Jr, Pinna D, Silacci C, Fernandez-Rodriguez BM, Vanzetta F, Santos C, Luke CJ, Torres-Velez FJ, Temperton NJ, Weiss RA, Sallusto F, Subbarao K, Lanzavecchia A.** 2010. Heterosubtypic neutralizing antibodies are produced by individuals immunized with a seasonal influenza vaccine. J Clin Invest **120:**1663–1673. http://dx.doi.org/10.1172/JCI41902.
15. **Ding H, Tsai C, Zhou F, Buchy P, Deubel V, Zhou P.** 2011. Heterosubtypic antibody response elicited with seasonal influenza vaccine corre-

lates partial protection against highly pathogenic H5N1 virus. PLoS One **6:**e17821. http://dx.doi.org/10.1371/journal.pone.0017821.

16. **Throsby M, van den Brink E, Jongeneelen M, Poon LLM, Alard P, Cornelissen L, Bakker A, Cox F, van Deventer E, Guan Y, Cinatl J, ter Meulen J, Lasters I, Carsetti R, Peiris M, de Kruif J, Goudsmit J.** 2008. Heterosubtypic neutralizing monoclonal antibodies cross-protective against H5N1 and H1N1 recovered from human IgM^+ memory B cells. PLoS One **3:**e3942. http://dx.doi.org/10.1371/journal.pone.0003942.
17. **Gerdil C.** 2003. The annual production cycle for influenza vaccine. Vaccine **21:**1776–1779. http://dx.doi.org/10.1016/S0264-410X(03)00071-9.
18. **Kelly J, Hurley D, Raghu G.** 2000. Comparison of the efficacy and cost effectiveness of pre-emptive therapy as directed by CMV antigenemia and prophylaxis with ganciclovir in lung transplant recipients. J Heart Lung Transplant **19:**355–359. http://dx.doi.org/10.1016/S1053-2498(00)00070-X.
19. **Hirji Z, O'Grady S, Bonham J, Mak M, Takata-Shewchuk J, Hawkins K, Gardam M, Law L, Mazzulli T, Conly J.** 2002. Utility of zanamivir for chemoprophylaxis of concomitant influenza A and B in a complex continuing care population. Infect Control Hosp Epidemiol **23:**604–608. http://dx.doi.org/10.1086/501979.
20. **Humar A, Siegal D, Moussa G, Kumar D.** 2005. A prospective assessment of valganciclovir for the treatment of cytomegalovirus infection and disease in transplant recipients. J Infect Dis **192:**1154–1157. http://dx.doi.org/10.1086/444398.
21. **Humar A, Kumar D, Preiksaitis J, Boivin G, Siegal D, Fenton J, Jackson K, Nia S, Lien D.** 2005. A trial of valganciclovir prophylaxis for cytomegalovirus prevention in lung transplant recipients. Am J Transplant **5:**1462–1468. http://dx.doi.org/10.1111/j.1600-6143.2005.00866.x.
22. **Yeo W, Hui EP, Chan AT, Ho WM, Lam KC, Chan PK, Mok TS, Lee JJ, Mo FK, Johnson PJ.** 2005. Prevention of hepatitis B virus reactivation in patients with nasopharyngeal carcinoma with lamivudine. Am J Clin Oncol **28:**379–384. http://dx.doi.org/10.1097/01.coc.0000159554.97885.88.
23. **Grigoleit GU, Kapp M, Hebart H, Fick K, Beck R, Jahn G, Einsele H.** 2007. Dendritic cell vaccination in allogeneic stem cell recipients: induction of human cytomegalovirus (HCMV)-specific cytotoxic T lymphocyte responses even in patients receiving a transplant from an HCMV-seronegative donor. J Infect Dis **196:**699–704. http://dx.doi.org/10.1086/520538.
24. **European Medicines Agency.** 2001. Note for guidance on choice of control group in clinical trials (CPMP/ICH/364/96). European Medicines Agency, London, United Kingdom. www.emea.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002925.pdf.
25. **Code of Federal Regulations.** 2014. Title 21. Food and drugs. Chapter I. Food and Drug Administration, Department of Health and Human Services. Subchapter D. Drugs for human use. Part 314. Applications for FDA approval to market a new drug. 21 CFR 314.126. www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=314.126.
26. **Presanis AM, Pebody RG, Paterson BJ, Tom BD, Birrell PJ, Charlett A, Lipsitch M, De Angelis D.** 2011. Changes in severity of 2009 pandemic A/H1N1 influenza in England: a Bayesian evidence synthesis. BMJ **343:**d5408. http://dx.doi.org/10.1136/bmj.d5408.
27. **McMichael AJ, Gotch FM, Dongworth DW, Clark A, Potter CW.** 1983. Declining T cell immunity to influenza, 1977–82. Lancet **2:**762–764. http://dx.doi.org/10.1016/S0140-6736(83)92297-3.