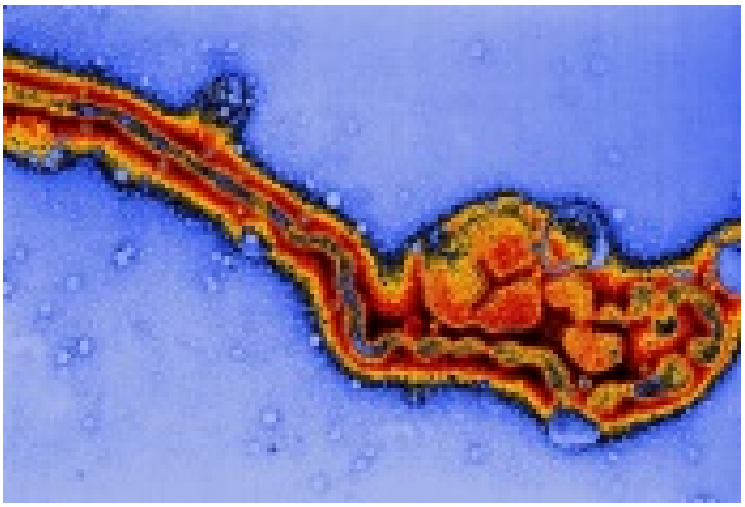
LANCASTER UNIVERSITY

Influenza C: a pilot study investigating the prevalence in Lancaster of a neglected respiratory virus



CDC / SCIENCE PHOTO LIBRARY / Universal Images Group

Rights Managed / For Education Use Only

Kate Atkinson

Submitted August 2015

This thesis is submitted in concordance with the requirements for the degree of Master of Science at Lancaster University

Declaration

I declare that this thesis is entirely my own work and has not been submitted elsewhere in another form for the award of a higher degree.

Acknowledgements

Firstly I would like to thank the patients and their families who consented to participate in this study.

Secondly I would like to thank my supervisors Dr Derek Gatherer (Lancaster University), Dr Mark Wilkinson (Royal Lancaster Infirmary), and Prof Roger Pickup (Lancaster University) who have supported and guided me throughout the duration of this project.

In addition I am grateful to Dr Lisa Bishop (Lancaster University), Dr Glenn Rhodes (CEH Lancaster), and Dr Robert Lauder (Lancaster University) who have helped me to learn the skills required for the laboratory work.

Dr Nicolas Salez (Aix-Marseille Université) has also been extremely helpful, and has provided various materials needed for the ELISAs and PCR such as positive controls and the influenza C antigen.

I also appreciate the help of the staff, both at the Royal Lancaster Infirmary and Queen Square Medical Practice, who have aided me in recruiting the participants for the study.

Finally I would like to thank the University Hospitals of Morecambe Bay Trust who have funded this intercalated MSc, and the Rosetrees Trust for the additional research grant.

Contents

List of Figures

List of Tables

Abbreviations

AH	Alpha haemolysin
BB	Blocking buffer
BEAST	Bayesian Evolutionary Analysis Sampling Trees
BEAUti	Bayesian Evolutionary Analysis Utility
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
CB	Carbonate buffer
cDNA	Complementary deoxyribonucleic acid
CM2	Matrix protein 2
CT	Cycle threshold
DNA	Deoxyribonucleic acid
ELISA	Enzyme-linked immunosorbent assay
EMM	Environmental Master Mix
Exp	Exponential clock
FluA	Influenza A
FluB	Influenza B
FluC	Influenza C
G	Relative centrifugal force
GTR	General time-reversible model
HA	Haemagglutinin glycoprotein

HE	Haemagglutinin-esterase glycoprotein
HEF	Haemagglutinin-esterase fusion glycoprotein
HI	Haemagglutination inhibition assays/titres
Ig	Immunoglobulin
IgA	Immunoglobulin A
IgG	Immunoglobulin G
HPD	Highest posterior densities
Log	Logarithmic clock
M1	Matrix protein 1
MEGA	Molecular Evolutionary Genetics Analysis
MRCA	Most recent common ancestor
NA	Neuraminidase glycoprotein
ND	NanoDrop
NEP	Nuclear export protein
NCBI	National Center for Biotechnology Information
NP	Nucleoprotein
NS	Non-structural protein
NS1	Non-structural protein 1
NS2	Non-structural protein 2
OD	Optical density
P3	Polymerase protein

PA	Polymerase protein
PB1	Polymerase protein 1
PB2	Polymerase protein 2
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline + tween 0.05%
PCR	Polymerase chain reaction
QSMP	Queen Square Medical Practice
R	Basic reproduction number
RLI	Royal Lancaster Infirmary
Rn	Fluorescence of the reporter dye divided by the fluorescence of the passive reference dye
ΔRn	The strength of the fluorescence signal at each time point during the PCR
RNA	Ribonucleic acid
SD	Standard deviation
S. aureus	*Staphylococcus aureus*
t-MRCA	Time of most recent common ancestor
TN93	Tamura-Nei, 93 model
vRNP	Viral ribonucleoprotein complex

Abstract

Influenza C is a virus found throughout the world that can cause respiratory illness, ranging from mild colds to pneumonias. Typically affecting younger children, it is a virus which can cause considerable illness and complications, and yet research into the virus is lacking.

This study aimed to identify the prevalence of influenza C in Lancaster, and to determine the current seropositivity levels in the study population.

148 participants were recruited to the study – 77 asymptomatic and 71 symptomatic. Participants were asked to provide a nasal or nasopharyngeal swab and/or a serum sample.

The swabs were analysed via polymerase chain reaction (PCR) and the serum samples by enzyme-linked immunosorbent assays (ELISA) and, of the samples tested via ELISA, 106 were positive (82%) and 23 were negative. Two swab samples also appeared positive for influenza C following PCR and so were sent for deep sequencing. A further seven mixed cDNA samples were also sent for deep sequencing to allow for comparison between different population groups etc.

Overall, it was found that influenza C is prevalent in the Lancaster area, with the entire study population having some level of exposure to the virus previously, although only 82% of participants met the threshold to be classed as seropositive, and two participants were actively carrying the live influenza C virus.

Further work needs to be done to analyse the seasonality of the virus and discover whether the virus has the same impact in the UK as it does in other parts of the world. As positive samples have been found and most of the population have influenza C antibodies, it provides a strong foundation for future work.

I Introduction

I.I Influenza C

Influenza C (fluC), first isolated in 1947 (1), is a virus that causes respiratory illness and infections. Despite being known about for over 60 years, fluC has been overlooked and research into this virus is currently lacking.

I.II *Orthomyxoviridae*

Influenza C is from the genus *Influenzavirus C* and family *Orthomyxoviridae* (2, 3), which also contains the other genera *Influenzavirus A* (fluA), *Influenzavirus B* (fluB), *Thogotovirus*, and *Isavirus* (3). Each virus has a genome composed of single-stranded RNA, varying from six (*Thogotovirus*) to eight (fluA, fluB, and *Isavirus*) segments (3). The influenza viruses are all commonly found in humans and some are also present in other species. *Thogotovirus* can also infect humans and livestock but is transmitted via ticks, whereas the fifth genus, *Isavirus,* is solely found in salmon and is known as infectious salmon anaemia virus (3). There has also recently been suggestion of an *Influenzavirus D*, but there is little research into this area (4).

I.III Influenza C Genome

FluC is an enveloped virus, and its genome comprises seven segments of single stranded RNA that is negatively polarised - PB2, PB1, P3, HE (or HEF), NP, MP, NS – numbered 1-7 respectively (5-7), and the ends of the RNA segments, 5’ and 3’, are complimentary (5). The segments code for nine proteins; three polymerase proteins (PB2, PB1, and P3), a haemagglutinin-esterase surface glycoprotein (HE) (see Figure 1), a nucleoprotein (NP), a matrix (M1) and a CM2 protein, and two non-structural proteins (NS1 and NS2/NEP) (5, 6)

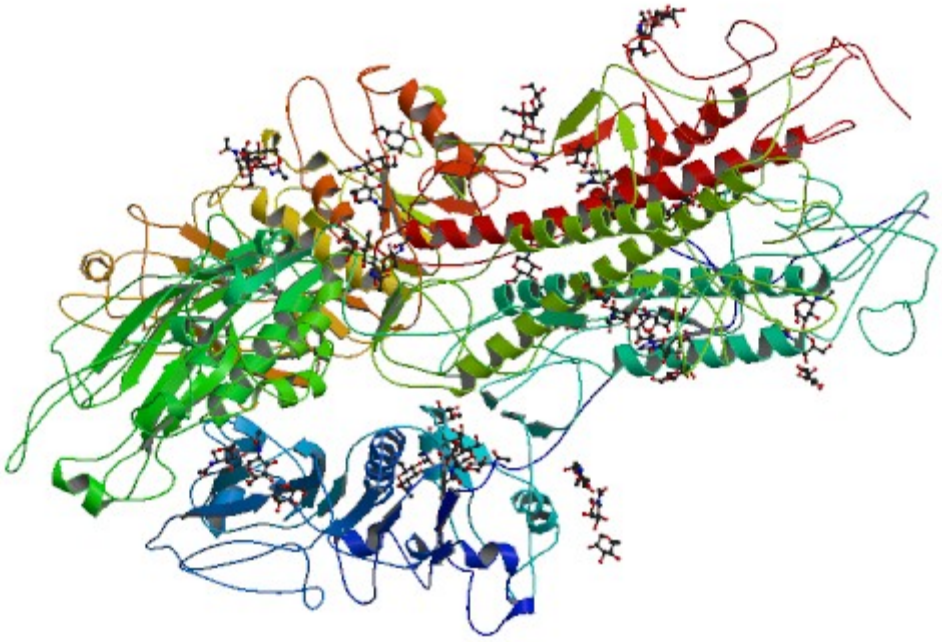


Figure 1 - X-ray structure of HE glycoprotein (Rosenthal *et al*. 1998, presented by PDB 2015) (9, 10)

The structure of the HE glycoprotein is comprised of many different protein chains, and each individual protein chain is coloured by position.

The role of each protein varies; PB2, PB1 and P3 are all part of the fluC RNA polymerase (6). HE takes part in receptor binding/destroying (acetylesterase) and membrane forming/fusion activities and forms a spike on the virion (7, 10-14). NP, when it encapsidates the RNA polymerase and RNA forms a viral ribonucleoprotein complex (vRNP) (15). M1 lies beneath the lipid bilayer and provides rigidity to the envelope, with CM2 also found in the membrane and provides a channel for chloride ions (16-18). NS1 aids viral mRNA splicing (19) whilst NS2 acts as a nuclear export factor (15) and is incorporated into fluC virions (20).

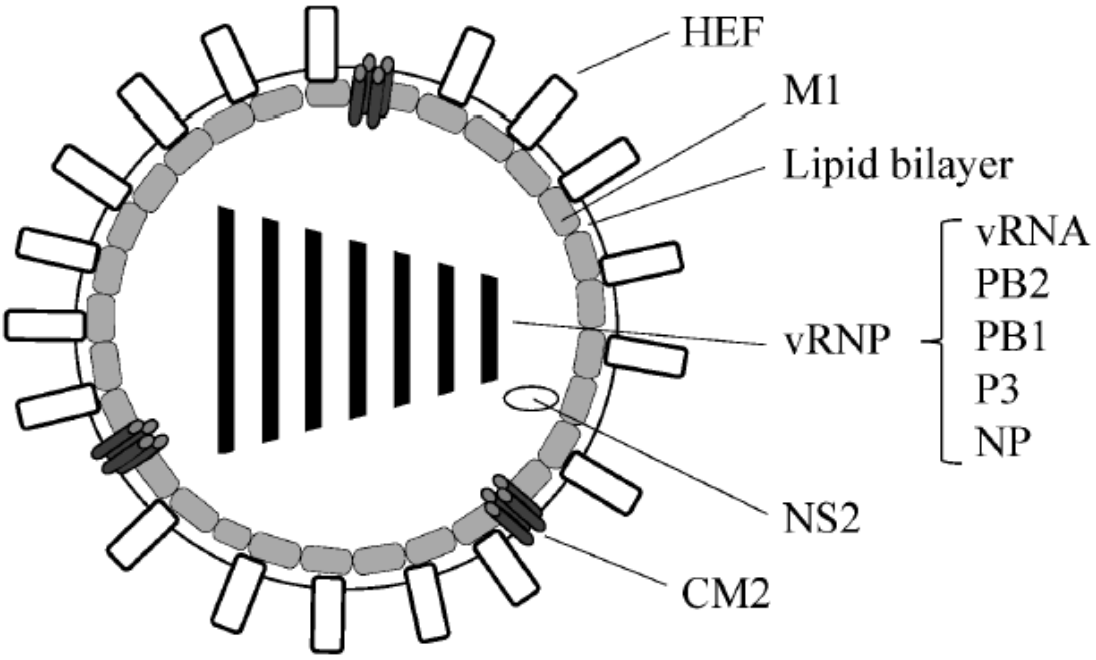


Figure 2 – Structure of fluC (Muraki and Hongo 2010) (7)

FluC is an enveloped virus, with a central viral ribonucleoprotein complex (vRNP) surrounded by a lipid bilayer. Eight of the nine fluC proteins are shown - three polymerase proteins (PB2, PB1, and P3), a haemagglutinin-esterase fusion protein (HEF), a nucleoprotein (NP), a matrix (M1) and a CM2 protein, and a non-structural proteins (NS2). The NP forms the vRNP by encapsidating the RNA polymerase (including PB2, PB1 and P3) and RNA, and the HEF forms a spike on the virion, M1 lies beneath the lipid bilayer and provides rigidity to the envelope, whilst CM2 provides a channel in the membrane for chloride ions and NS2 acts as a nuclear export factor.

I.IV Comparison of the influenza viruses

I.IV.I Genomes

Of the three influenza viruses, fluA and fluB are more closely related, with fluC the outlier. FluA and fluB both have eight RNA segments as, instead of the HE segment present on fluC, fluA and fluB have two separate segments – haemagglutinin (HA) and neuraminidase (NA)

(3, 6). Another difference between these segments is that the cytoplasmic section of the HE segment on fluC contains three amino acids, whereas the HA segment contains 11 to 12 (7). The P3 segment of fluC correlates to the PA segment on fluA and fluB, and is so named as when it is at a neutral pH, it does not display any acid charge features (7).

In terms of the host range of fluC, it appears to be more closely related to fluB than it is to fluA. FluB is generally only found in humans, although it has been reported in seals (21, 22), and fluC is mainly found in human and swine species (3), whereas fluA, in contrast, is found in various hosts including humans, swine, horses, and birds (3). The large avian fluA reservoir is nearly always the cause of the fluA pandemics (except 2009), due to the vast number of strains capable of reassortment. As fluB does not have a known animal reservoir, like fluA, it is less capable of producing a pandemic (23), and interestingly, despite fluB being more 'well known' and better documented than fluC, some studies have found more patients infected with fluC than fluB (24, 25).

I.IV.II Seasonality

Opinions on the seasonality of fluC seem to differ, and this may be due to the different regions and conditions where fluC is found throughout the world. One view is that unlike fluA, which appears to be seasonal and causes epidemics/pandemics (26-28), fluC and fluB are present in the human population all year round (3), which appears to be the case in some parts of Spain (25) and France (29). It does appear that there has previously been a national fluC epidemic in Japan in 2004 (30), and fluC often coincides with fluA and fluB infections (25, 28), but otherwise fluC is believed to be present in the population continuously. The other opinion is that fluC is seasonal and is found in the winter to early summer months, particularly January to June, which is the case in Japan and the Philippines (27, 30-33), Spain (24), Finland (28), Canada (34), France (35) , and Nigeria (36).

I.V Symptoms of influenza C infection

Common symptoms of fluC are fever, cough and rhinorrhoea (33, 37), with 33% of adults experiencing fever (38) and 90% of children (37). On average, symptoms last for six days (range 5-11 days), with a mean incubation period of four days (range 3-6 days) (38). It is also worth noting that infection with fluC, like fluA and fluB, can cause some economic damage, due to absence from work etc. (39) and so provides another reason for studying this topic.

The illnesses attributed to fluC are generally mild, upper respiratory tract infections in adults and adolescents (38, 40, 41), however in young children fluC can also cause lower respiratory tract infections, including bronchitis, bronchiolitis and pneumonia, with a high risk of complications, particularly in children under two years of age (24, 30, 33-35, 37, 42-46). One case documented that fluC caused an episode of acute encephalopathy (47) and in younger children, fluC also appears to be a common cause of gastrointestinal symptoms (35, 37).

It has also been found that fluC is more prevalent in children under the age of five years (35, 48), although elderly patients have also been found to have fluC infections (35, 49). Homma *et al.* also suggested that primary fluC infections occur in those under seven years old (50), and so it is important that swabs from children are included in this study.

I.V.I Transmission and replication

In humans, fluC is spread via respiratory secretions; when the infected host coughs or sneezes the virus becomes airborne and is transmitted to another individual (11, 23). The virus then replicates in the trachea-bronchial epithelial cells (23), which is typical for influenza in human hosts. The activated HA/HE segment of the virus binds to sialic acid, which is present on host cell glycoproteins on the plasma membrane (3, 23). The entire replication of the RNA takes place in the nucleus of the cell (3). In contrast, avian fluA is spread via the faecal-oral

route (3, 11, 23). Bird species, such as ducks have receptors for the virus in their gut (3, 23), and it has also been documented that humans have been infected with avian strains of fluA in their gut (51), but there is currently no evidence for such receptors in fluC (35).

I.VI Immune response

I.VI.I Immunoglobulin G (IgG)

Immunoglobulin G (IgG) is one of the five classes of Ig expressed by B cells (52). The IgG antigen receptor is comprised of two heavy polypeptide chains and two light polypeptide chains, linked by disulphide bridges (found in the constant region). Most of the chain remains constant between each Ig class; however the heavy-chain constant region differs between the four Ig classes (52). The IgG antibody is soluble and can be found in all bodily fluids, including blood, tears, saliva and breast milk (52).

Interestingly, following intranasal or subcutaneous vaccination with fluA, the highest level of antibody found in nasal secretions was IgA. This IgA was different to the IgA found in serum (the nasal IgA contains extra-antigenic determinants, which differentiates it from the IgA found in serum), but both the serum IgA and IgG were also found in reasonable quantities in nasal secretions (53). In comparison, the highest level of antibody found in serum following vaccination was IgG. It was also of note that children were more responsive than adults to intranasal vaccination, and generated higher levels of antibodies (52).

Table 1 shows a range of normal IgG antibody levels for all ages. These values were calculated using nephelometry, and will be used as the normal range reference values for the total IgG antibody levels. The range of values for adults is 639-1349 mg/dl (53, 54), which is similar to values in other articles (694-1618 mg/dl total IgG (55), 584-1486 mg/dl (56)). Other methods have been used to calculate IgG levels, including ELISA, radial immunodiffusion and immunoturbidimetry; one paper calculated the IgG reference values

using immunoturbidimetry, and found that the mean (1058 mg/dl) and median (1073 mg/dl) fell within those values listed (56). Values for IgG deficiency were also included, with deficiency defined as <700 mg/dl total IgG (55), or below the 95% confidence interval (584 mg/dl) (56), however it was shown that those donors who had values that would class them as 'deficient' were actually healthy, and so any conclusions drawn need to be done so carefully (56).

Table 1 - Normal values - total IgG antibody concentration (Joliff *et al*. 1982, adapted Agarwal *et al.* 2007) (53, 54)

This shows the mean total IgG concentration expected for people in different age groups, and the normal range. These values will provide the basis for the expected total IgG levels for participants in this study.

Age	IgG, mean (range), mg/dl
Cord blood (term)	1,121 (636–1606)
1 mo	503 (251–906)
2 mo	365 (206–601)
3 mo	334 (176–581)
4 mo	343 (196–558)
5 mo	403 (172–814)
6 mo	407 (215–704)
7–9 mo	475 (217–904)
10–12 mo	594 (294–1,069)
1 y	679 (345–1,213)
2 y	685 (424–1,051)
3 y	728 (441–1,135)
4–5 y	780 (463–1,236)
6–8 y	915 (633–1,280)
9–10 y	1,007 (608–1,572)
Adult	994 (639–1,349)

Table 2 – Normal values - fluA IgG antibody concentration (Crum-Cianflone *et al.* 2012) (55)

This table shows the normal mean and range fluA IgG levels. These values will be used to provide an estimation of the fluC IgG values expected to be found in this study.

HIV-uninfected adults			
Factor	Total group *n* = 24	Robust antibody response *n* = 12	Poor antibody response *n* = 12
Baseline overall IgG levels, mg/dl			
Median Total IgG	1050 (890–1150)	1060 (966–1130)	933 (865–1180)
Baseline influenza-specific IgG levels, median levels, mg/dl			
Total fluA IgG	24 (15-39)	21 (7-30)	24 (15-40)
IgG1	17 (13-35)	15 (5-27)	22 (14-35)
IgG2	1 (0-2)	1 (0-2)	1 (0-2)
IgG3	3 (0-6)	1 (0-5)	3 (1-6)
IgG4	0 (0-0)	0 (0-1)	0 (0-0)

I.VI.II IgG in response to the influenza virus

Figure 3 shows a cell infected with the influenza virus. The infected cell expresses antigens on the surface which the IgG binds to. Effector cells then bind to certain regions on the IgG antibodies, releasing cytotoxic granules and antiviral cytokines which ultimately results in death (apoptosis) of the infected cell (57).

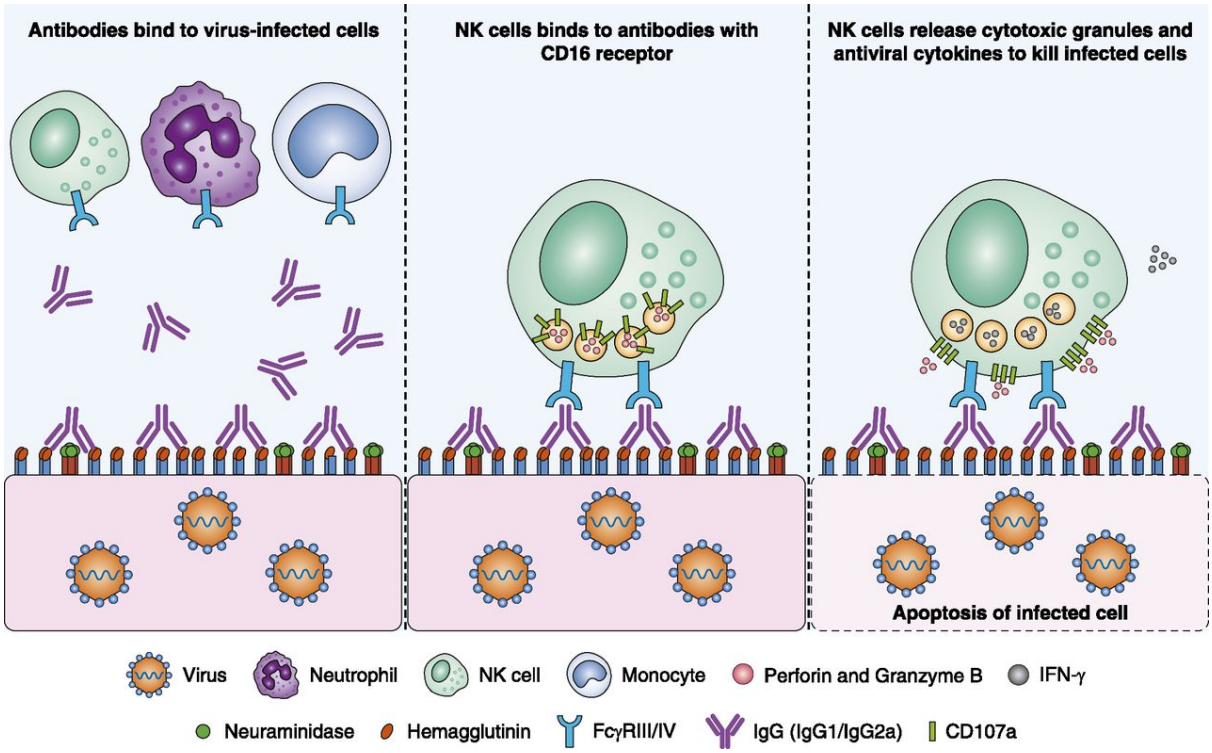


Figure 3 – IgG response to influenza infected cells (Jegaskanda *et al.* 2014) (57)

The human immune response to an influenza infection. Specific IgG antibodies against the specific influenza virus bind to infected cells. Natural killer cells then bind to the IgG antibodies which have the CD16 receptor. Perforin and granzyme B are then released by the natural killer cells, which are programmed to kill the infected cells.

I.VII Antigenic variation

I.VII.I Reassortment

Reassortment occurs within human hosts when they are infected with two different strains of fluC via exchange of segments (31, 32, 58-63), and this also occurs in fluB infections (64).

I.VII.II Antigenic drift

All three influenza viruses are capable of antigenic drift, although it occurs at a faster rate in fluA and fluB (11, 15, 65). This is a process in which mutations can occur naturally in one or

two glycoproteins – HA and NA (15). These proteins are surface antigens on the virus and have slightly differing roles; HA binds to receptors on host cells whereas NA is an enzyme that removes the sialic acid that HA binds to, therefore freeing the virus particles from the host cell receptors, allowing virions to escape from the cell, and aiding the spread of the virus (10, 15, 22). The mutations occur within the antibody-binding sites in HA/NA and accumulate over time. As these mutations occur, the antibodies that the host produced can no longer inhibit the virus because the structure of the HA and NA has changed and no longer produces a match for the antibody and so the virus can therefore spread more easily (15). It is this process that is responsible for the annual epidemics and seasonality seen in human fluA and fluB infections (22). The process is slightly slower in fluC, due to the presence of the HE glycoprotein instead, but nevertheless it has been noted that the dominant antigenic group on fluC is replaced every several years (26), showing that the process does occur.

By contrast, some papers have reported that antigenic drift does not occur in fluC, and instead antigenic replacement occurs (due to immune selection in those who have cross reactive antibodies) (25, 26). Interestingly, fluA is also capable of antigenic shift (15, 22), which can be seen as analogous to the process of antigenic replacement in fluC, in that both produce fairly sudden changes in antigenicity.

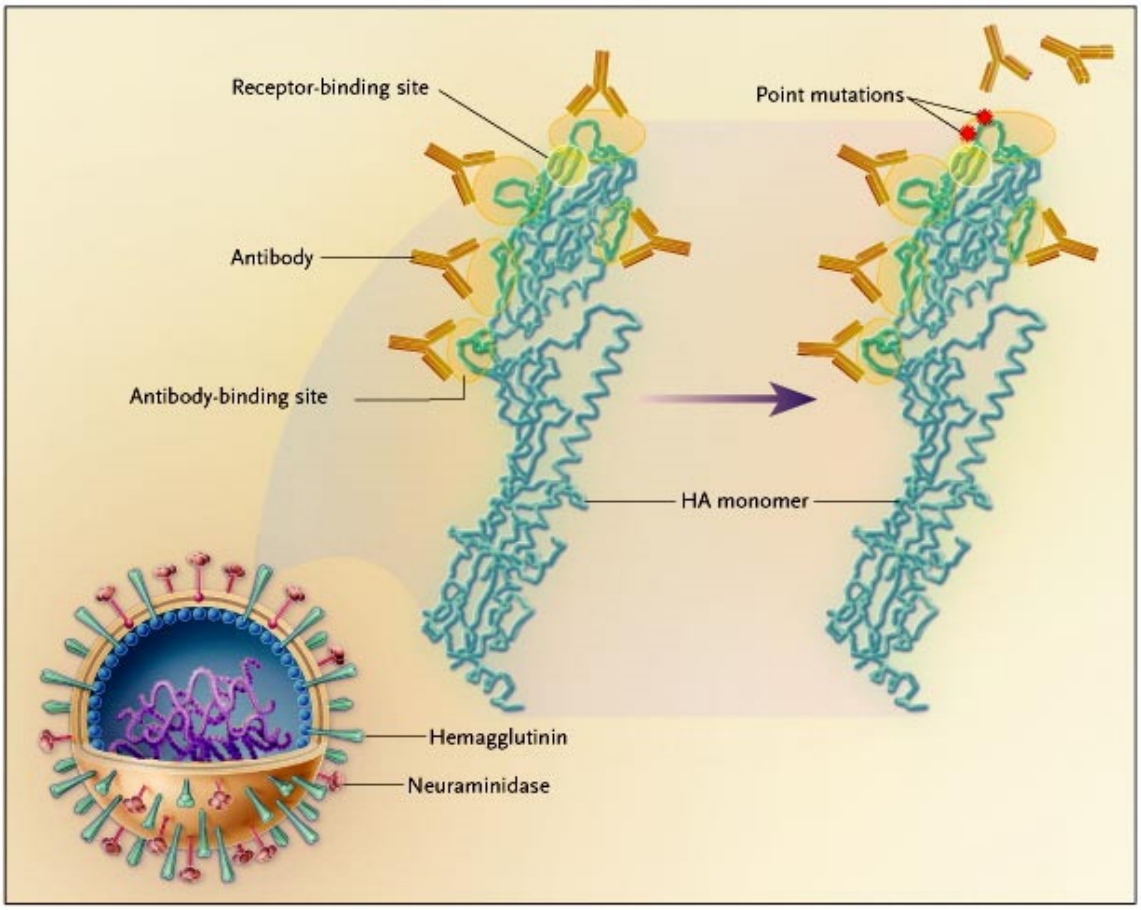


Figure 4 - Structure of a HA Monomer and Location of the Antibody-Binding Sites (Treanor 2004) (66)

The haemagglutinin (HA) and neuraminidase (NA) glycoproteins on fluA are shown. When antigenic drift occurs, point mutations at the antibody-binding sites mean that the antibodies to HA/NA can no longer bind (15).

I.VII.III Rate of nucleotide substitution

Another aspect of antigenic variation is nucleotide substitutions, and the rate at which this occurs. Each segment has a different rate of nucleotide substitution, and this also differs between fluA, fluB and fluC, for example the rate of nucleotide substitution of the fluC HE segment is one ninth of the speed of the equivalent human HA segment on fluA ($4.3x10^{-3}$

[H1] and $4.44x10^{-3}$ [H3] (15)) (6, 7). Table 3 shows the nucleotide substitution rates for all seven of the fluC segments. The rate of the substitutions is x 10^{-4} per site per year. The time of most recent common ancestor (t-MRCA) is also shown, along with the 95% highest posterior densities (HPD) for each segment. Pi represents the average number of substitutions per site in the alignment.

Table 3 – Nucleotide substitution rates for the fluC segments (Gatherer 2010) (58)

This table shows the nucleotide substitution rates (x 10^{-4} per site per year) for the seven different fluC segments. Pi represents the average number of substitutions per site in the alignment. The time of the most recent common ancestor and the 95% highest posterior densities are also shown.

Segment	Gene	Pi	rate	95% HPD	t-MRCA	95% HPD
2	PB1	0.027	6.77	4.78-8.89	1944	1940-1948
5	NP	0.020	5.74	4.05-7.56	1938	1924-1951
3	P3	0.026	5.49	2.94-8.25	1938	1920-1949
4	HE	0.042	4.87	4.11-5.66	1890	1857-1924
1	PB2	0.017	4.32	2.90-5.86	1936	1922-1947
6	M	0.021	4.12	3.29-4.99	1937	1930-1944
7	NS	0.020	3.72	2.70-4.80	1916	1891-1937

I.VII.IV Vaccine

Due to the capabilities of the viruses to undergo antigenic variation, creating a vaccine is difficult. A new vaccine is produced every year, against fluA and fluB, and the epidemic activity of each new virus has to be predicted, in order to select which strains will be included (15). As fluC does not undergo changes as frequently as fluA and fluB, and in older articles it was thought to be less of a threat to humans, it is not currently deemed important enough to be included in the vaccine.

I.VIII Seroprevalence

In the studies previously carried out, fluC global seroprevalence appears to be over 50% (28, 44, 67-69), with values ranging from 50% in Reunion Island (28) to 89% in Finland (27), although one article reports a seroprevalence rate of around 100% in those aged over 10 (46). With regards to the UK, seroprevalence values have been reported between 60-70% for those aged over 25 years (28, 70). These slight variances throughout the world could be due to a number of reasons, for example it could be due to different population groups around the world, or there could be methodological differences such as different detection methods, different sample types, different sample preparation methods or different sources (people).

It is currently accepted that antibodies against fluC are acquired early on in life, with levels peaking during early adulthood (42, 44, 49, 71, 72). It has also been noted that these antibodies can then drop away again over the age of 65 years, as shown by O'Callaghan *et al.*, who found that the incidence in those over 65 was 18.3%, compared to 96% in young adults (68). However, despite developing these antibodies it appears that it is possible to be re-infected with fluC (37, 42, 49, 71), which supports the statement that fluC is capable of antigenic variation and can therefore reinfect the same host multiple times.

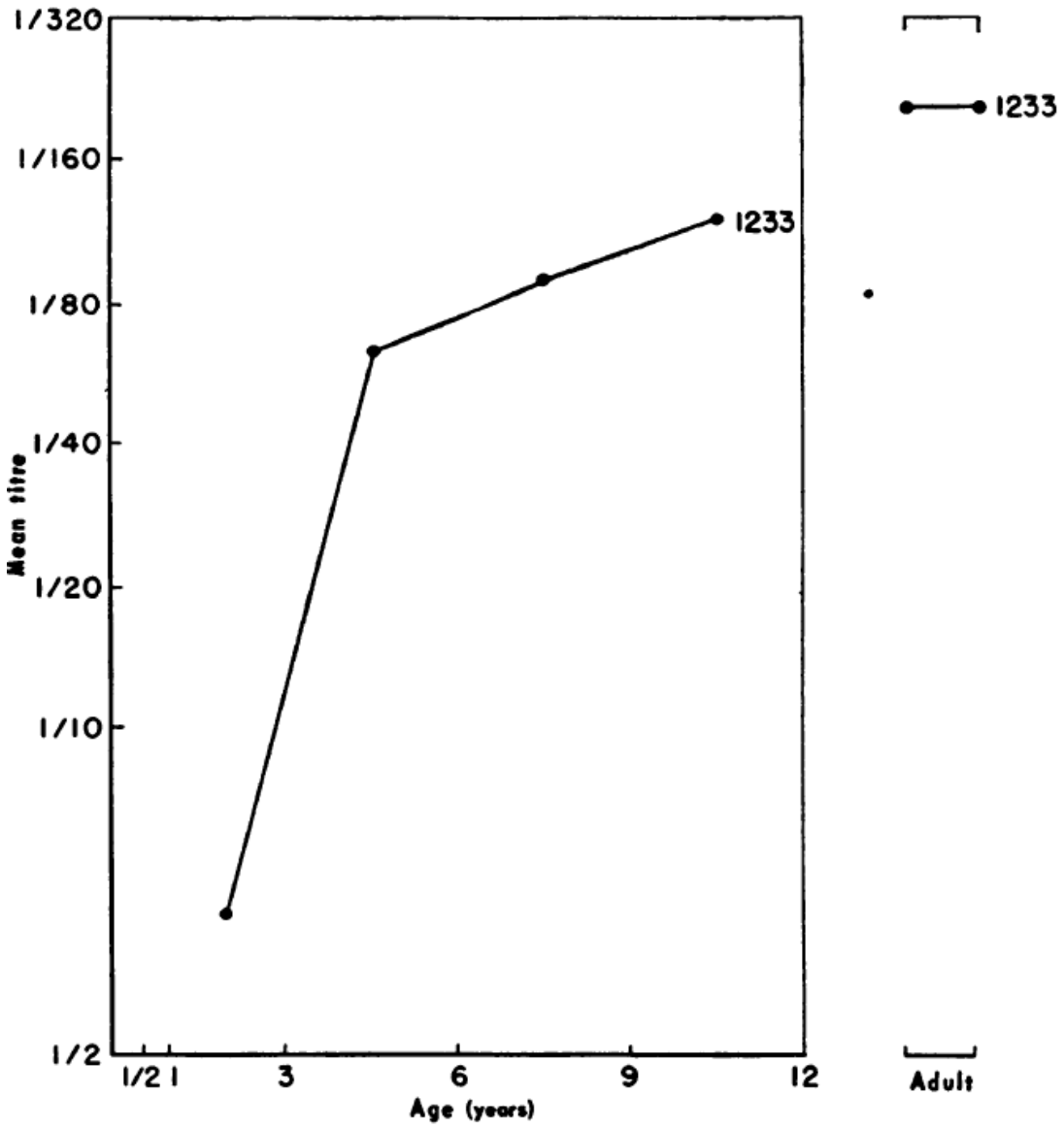


Figure 5 – Mean fluC sera antibody titres (Hilleman *et al.* 1953) (72)
The mean antibody titres against fluC (1233) increase with age, with the levels plateauing in early adulthood. The levels then continue at a steady state.

I.IX Epidemiology of fluC

I.IX.I Nomenclature

FluC viruses are named based upon a series of information; firstly the influenza virus type (C), then the host e.g. swine (although this is not included if the virus is isolated from

humans), then the geographical location, the allocated strain number, and finally the year of isolation.

In contrast, fluA is currently classified into different types according to the HA and NA segments, and the subtypes (HA or NA) are added at the end (3, 22). There are 16 HA and nine NA types (3), all of which are found in avian species (10).

I.IX.II Lineages

Different genes are found to co-circulate in the population at any one time (73). Antigenic and genetic analysis of the HE segment of various fluC strains has shown that they can be split into six different lineages; Taylor/47, Kanagawa/1/76 (KA176), Yamagata/26/81 (YA2681), Aichi/1/81 (AI181), Sao Paulo/378/82 (SP82), and Mississippi/80 (MS80) (6, 7, 31). (Figure 6 shows these six different lineages, presented in a phylogenetic tree). The HE segment has also been classified using the alternative method of Bayesian statistics (58). The most recent common ancestor (MRCA) of the HE gene can be traced back to 1890, whereas for PB1 the most recent common ancestor is in 1944 (58) (Table 3).

The most recent data on genome constellations is from Japan in the 1990s (58), and the most recent full genome was isolated in 2011 (74), and so any further changes in the fluC lineage would be near impossible to plot as the data are lacking.

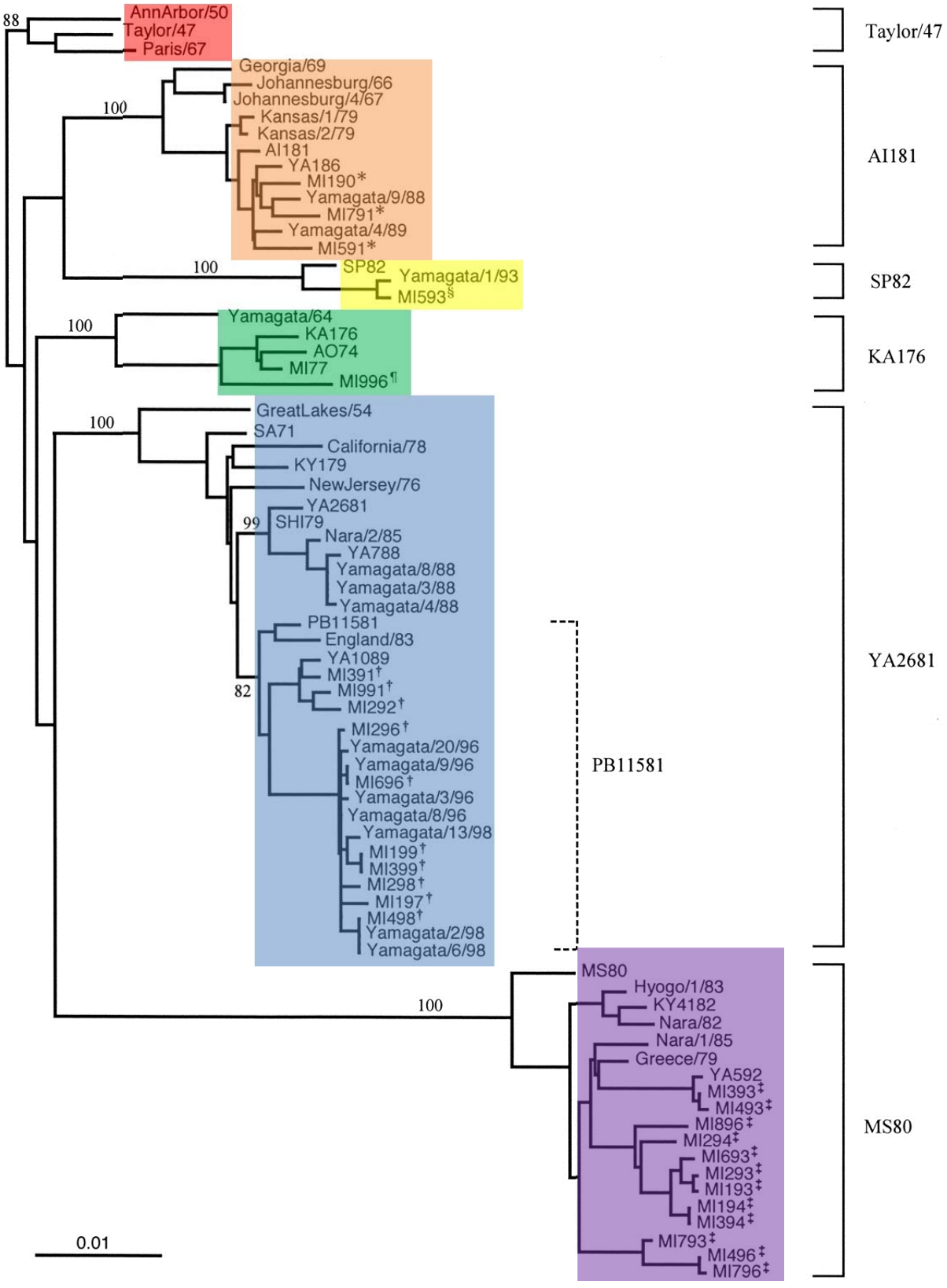


Figure 6 - Phylogenetic tree of fluC virus HE genes (Matsuzaki *et al.* 2003) (31)

The different strains can be categorised into one of six lineages - Taylor/47 (red), Kanagawa/1/76 (KA176) (green), Yamagata/26/81 (YA2681) (blue), Aichi/1/81 (AI181) (orange), Sao Paulo/378/82 (SP82) (yellow), and Mississippi/80 (MS80) (purple).

I.X Studies testing serum for fluC antibodies

A few studies have tested serum samples for antibodies against fluC.

Haemagglutination inhibition assays/titres (HI) have been commonly studied, as most adult serum samples contain haemagglutination-inhibiting antibodies (25, 26, 28, 32, 44, 49, 61, 67-69, 72). Despite this, HI is now quite an old method and newer methods, such as ELISA, have been shown to be simpler (75-77).

ELISAs have also been used in studies to detect fluC antibodies (27, 28). One of the most recently published papers by Salez *et al.* performed both an HI and an ELISA on serum samples and found very similar results, with the ELISA having a high positive predictive value (91%) and specificity (86%), and modest sensitivity (66%) (28). One paper using an ELISA to detect fluA antibodies also created standard curves using known values of IgG, which allowed for quantification of IgG antibody levels in each sample (55).

As described by Homma *et al.* newborns under the age of two months have maternal antibodies against fluC (49). As these antibodies wear off children become more susceptible to fluC infection and are likely to become infected before they develop their own antibodies. The antibody titres then increase and peak during adolescence/early adulthood (49, 72). Due to this trend, it is important to try and collect samples from patients of all age groups in order to try and monitor this trend.

I.X.I Capillary blood sampling

Although capillary blood samples provide smaller volumes they are less painful and invasive then venous blood samples, making them more appealing to children, and can give similar ELISA results to venous blood samples, when looking for antibodies against rubella (75-77). Figure 7 shows the results of Vaheri *et al.* when comparing venous and capillary blood samples – the correlation coefficients between venous serum and finger-tip blood and venous serum and finger-tip plasma are (r) = 0.945 and (r) = 0.949 respectively (76). Vejtorp *et al.* also recommended using BD Microtainer® tubes to store the capillary blood samples prior to performing the ELISA, as they are simple to use and store and the samples do not degrade as they can do when collected on filter paper (75).

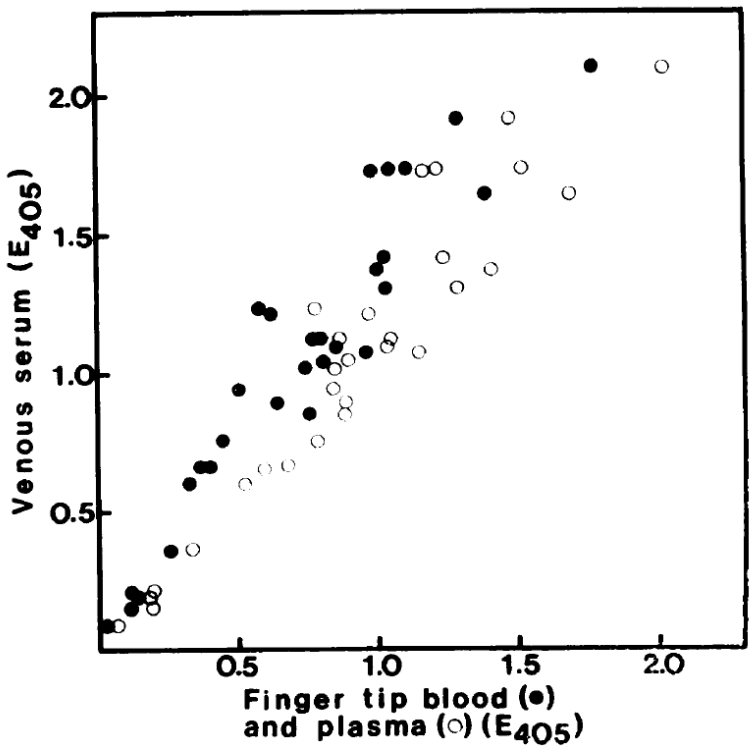


Figure 7 – Comparison of venous and finger-tip (capillary) blood samples when performing an ELISA to detect rubella antibody levels (Vaheri *et al.* 1980) (76)

Venous and capillary blood samples were compared, and the correlation coefficients between venous serum and finger-tip blood and venous serum and finger-tip plasma are (r) = 0.945 and (r) = 0.949 respectively. Both sets of results give very similar readings.

I.XI Studies testing samples for fluC

When looking for fluC during an active infection, nasal swabs (24, 28, 45, 47, 74), pharyngeal swabs (44), nasopharyngeal swabs (26, 32), oropharyngeal swabs (24), throat swabs (32, 47, 61, 74), nasal aspirates (44), nasopharyngeal aspirates (23, 78) and sputum samples (27) have been used. The general method used for analysing these was PCR (24, 27, 32, 47), with only the targeted gene-specific primers varying (HE - (26, 44, 74), MP - (28), NS - (45), NP - (44)). A recent paper, by Salez *et al.* used nasal swabs from patients of all ages with acute respiratory symptoms to check for the presence of fluC (28). This was done using a TaqMan® real time RT-PCR looking for primers and probes on the matrix gene (28).

There is, however, some discrepancy as to whether to use nasopharyngeal swabs or sputum samples to gain samples for detection of active fluC infections. Some papers state that sputum samples are better (79, 80), but not every patient can expectorate (particularly healthy participants) which would therefore suggest that nasopharyngeal swabs would be better, as every participant could provide a sample if they gave consent. It was also found in one of these papers that there was no significant difference between nasopharyngeal swab and sputum detection rates for fluA and fluB in culture or immunoassay (culture ($P = 0.15$), immunoassay ($P = 0.07$)) (79), which would also support the use of nasopharyngeal swabs over sputum samples. One study, by Kauppila *et al.* stated that despite analysing sputum samples, for future research they would recommend using an alternative such as nasal swabs or nasopharyngeal aspirates as in some cases they described it as “impossible” to collect sputum samples (27).

I.XII *Staphylococcus aureus*

By analysing the swabs for the presence of *S. aureus*, it will prove that the swabs have made contact with the lining of the nares or nasopharynx and therefore that material has been collected on the swabs.

Staphylococcus aureus (*S. aureus*) is a Gram-positive coccus from the genus *Staphylococcus. Staphylococci* were traditionally divided into two groups – those that were coagulase-positive and those that were coagulase-negative. The coagulase-positive *Staphylococcus* group, which is able to clot blood plasma (via the coagulase reaction), is only comprised of *S. aureus* and is thought to be more pathogenic than other *Staphylococci* (81), whereas the coagulase-negative *Staphylococci* are comprised of over 30 species (81).

S. aureus, part of the natural flora, colonises the skin of humans and is most commonly found in the anterior nares (82). People can be categorised into one of three groups: persistent carriers (approx. 20%), intermittent carriers (around 30%) and non-carriers (about 50%) (82). Criteria for classification into one of these groups varies between studies, but it can generally be accepted that a person would have all positive nasal swabs to be classed as 'persistent', at least one positive swab to be classed as 'intermediate' and no positive swabs to be classed as a non-carrier (82). General surveys of the population suggest nasal carriage rates of 35%, since 1934, (83) and of 27%, since 2000 (82), which suggests a decrease in nasal carriage rates over time. Higher carriage rates were also found in men (84, 85) and newborn babies (86), with levels decreasing into adulthood (87). There is also a strong correlation between nasal carriage and hand carriage (88) – with hands being the main vector for transmission of *S.aureus* e.g. by nose picking (89).

I.XIII Deep Sequencing

Deep sequencing is a process that allows the full genome of a virus, or similar, to be identified. Any nucleotide sequences in a given sample are detected, depending upon the quality of the sample provided and the specific technique used (90). In the case of fluC, a high quality sample would enable the seven RNA segments to be isolated, and these data can then be analysed via metagenomics, the study of genomes collected from samples. This allows for greater understanding of the composition and behaviour of the viruses. For example, it allows for comparison between different strains of the virus and predictions as to which strains have evolved from others, as shown in the six different lineages in Figure 6. Phan *et al*. used metagenomics in the analysis of cyclovirus found in paediatric nasopharyngeal aspirates, which showed how metagenomics can be used to identify nucleic acids and indicate which viruses they are more closely related to (91).

I.XIV Summary

Influenza C is a respiratory virus which is currently under-researched and not well-documented. Recent papers have suggested that the effects of infection with this virus are more severe than initially reported, particularly in younger children (46).

There is a growing need for the importance of this virus to be outlined and its clinical significance to be described in further detail. Investigation of both patients and the general population will allow for further analysis of the fluC virus.

Serum samples and nasal/nasopharyngeal swabs will show seropositivity levels and active fluC infections, respectively.

I.XV Aims

This pilot study aims to identify the prevalence of fluC in the Lancaster area and also whether patients have previously been exposed to the fluC virus and therefore have antibodies against it.

In order to achieve this, the following criteria will need to be met:

- Nasopharyngeal and serum samples from patients of all ages with respiratory or 'flu-like' symptoms will need to be collected, along with samples from asymptomatic participants.
- A robust method for polymerase chain reaction **(**PCR) will need to be developed in order to process the nasopharyngeal swabs.
- An appropriate enzyme-linked immunosorbent assay (ELISA) protocol will need to be produced to enable the serum samples to be analysed.
- Any samples which are positive for fluC or may be deemed interesting will be sent for further deep sequencing at the Institute of Biological, Environmental and Rural Sciences (IBERS) Phenomics Centre, Aberystwyth University.

II Materials and Methods

II.I Materials

II.I.I Polymerase chain reaction (PCR) materials

- Nasopharyngeal swabs (MW951SENT, Sigma Virocult® ENT, Medical Wire) / Nasal swabs (Copan)
- MagMAX™ Viral RNA Isolation Kit (Ambion)
- High-Capacity RNA-to-cDNA™ Kit (Applied Biosystems®, Life Technologies™)
- TaqMan® Environmental Master Mix 2.0 (Applied Biosystems®, Life Technologies™)
- Positive control for qPCR (from Dr Nicolas Salez (Aix-Marseille Université))
- Primers and probe sequences for the matrix gene:
    - Forward: CAT AAT TGA ACT TGT CAA TGG TTT TGT
    - Reverse: TTC AGG CAT AAT TGT GGT CTT TAT ATC T
    - Probe: FAM-CTC GGC AGA TGG GAG AGA TGG TGT GTAMRA (28)

II.I.II Enzyme-linked immunosorbent assay (ELISA) materials

- Serum samples
    - Venous samples collected using a red-topped BD Vacutainer® serum tube 10ml, Silica (Clot Activator)
    - Capillary samples collected using gold-topped BD Microtainer® Serum Separator Tubes (SST™) with added polymer gel
- FluC Antigen, positive sera and negative sera (provided by Nicolas Salez)

II.II Ethical approval

This study received a favourable opinion from NRES Committee London - Queen Square, reference: 14/LO/1634.

The study has also been approved to be a part of the NIHR Clinical Research Network Portfolio: IRAS Project 147631, UKCRN ID 17799, Influenza C: a pilot study on a neglected respiratory virus.

II.III Patient recruitment and sample collection

Patients of all ages were recruited from the Royal Lancaster Infirmary (RLI), Lancaster, and Queen Square Medical Practice (QSMP), Lancaster. Additional samples were also provided by participants at Lancaster University. Seventy one samples were taken from patients with respiratory symptoms and a further 77 samples were taken from healthy 'asymptomatic' participants.

II.III.I Inclusion criteria

Any participants who were currently suffering from 'Influenza like illness' (defined as presence of one of the following: sudden onset of fever, nasal discharge, sore throat/cough, fatigue/body ache, difficulty in breathing (74)) were included in the symptomatic group.

All patients included in the study were allocated a study number ranging from K001 – K148.

II.III.II Consent

Consent was provided by all participants with capacity over the age of 16. For those participants under the age of 16, parental consent was given in addition to the assent of the child. Any adult participants who lacked capacity for a transient period of time had a named consultee to give consent on their behalf, with the patient themselves asked to give their

consent in retrospect following regaining their decision-making capacity. If the patrticipant did not regain the ability to consent, then the consultee's consent was taken on their behalf.

II.III.III Samples

Each participant had the option to consent to one, all, or none of the following samples:

- A serum sample - for those over the age of 16.
- A capillary blood sample – a heel prick for infants and a finger prick for older children (or adults, if venous sampling equipment wasn't available).
- A nasopharyngeal or nasal swab, depending upon the availability of swabs.

Sample collection commenced in November 2014 and was completed in May 2015. A total of 148 samples were collected in this time; 71 symptomatic participants and 77 asymptomatic.

II.III.IV Serum sample preparation

From all consenting adult participants, 10ml of venous blood was extracted using a red-topped BD Vacutainer® tube. The blood samples were then left at room temperature to clot for around 15-30 minutes. The samples were then placed on ice for transportation to the laboratory where they were then centrifuged for 10 minutes at 1000-2000g (relative centrifugal force (g)) and 4°C. Once separated, the serum was extracted from the tube and placed into multiple anonymised Eppendorf tubes. Following this, the samples were then stored at -80^{o}C until they were to be analysed.

II.III.V Capillary sample preparation

Patients under the age of 16 years, or adults who did not wish to provide a venous blood sample were able to provide a capillary blood sample instead. These samples were collected in gold-topped BD Microtainer® Serum Separator Tubes (SST™) with added polymer gel.

Once collected, the samples were allowed to clot for 30 minutes before being transported on ice to the laboratory. Following this, the tubes were placed into a centrifuge for at least 90 seconds at 6000-15000g. The polymer barrier material within the tube then lies between the serum and blood cells, allowing the serum to be extracted as required. Samples were then placed in a freezer at-80^{o}C until they were analysed.

II.IV PCR

II.IV.I RNA Isolation

As the nasal swabs did not come with a preservative, an additional step had to be performed on the nasal swabs. The swab tips were immersed in 500µl PBS, and then were vortexed for around two minutes in order to mix the samples into the solution. The swab tips were then left suspended in the PBS until they were ready to be used for the RNA extraction.

The MagMAX™ Viral RNA Isolation Kit (AM1939, Ambion®, Applied Biosystems™, Life Technologies™) was used to extract viral nucleic acids from all of the nasal/nasopharyngeal swab samples. The instructions provided by Applied Biosystems™ were followed, with 400µl of sample used, resulting in a 50µl final volume, all of which were prepared in sterile 1.5ml Eppendorf tubes. The carrier RNA that was supplied was replaced with linear acrylamide (5 mg/ml) (AM9520, Ambion®, Invitrogen™, Life Technologies™), in order to prevent the carrier RNA from masking any nucleotides in the samples. The same process was also followed using 400µl from an unused sterile swab, in order to act as a negative control.

II.IV.II NanoDrop

The quality and quantity of RNA extracted from samples was assessed by spectrophotometry using the NanoDrop® 1000 Spectrophotometer V3.3.0 (Thermo Fisher Scientific). The NanoDrop machine can generally accurately detect samples with a concentration above

2ng/µl, and so any sample with a value greater than two would be classed as having RNA. 2µl of each sample was used in the spectrophotometer, and the remaining 48µl was then stored at -20^{o}C.

II.IV.III RNA to cDNA

RNA was converted to cDNA using the High Capacity RNA-to-cDNA Kit (Applied Biosystems®, Life Technologies™), following the manufacturer's instructions. Briefly, these were: 20.0µl reactions (9µl sample (RNA), 10µl 2X RT buffer, 1µl 20X RT enzyme mix) prepared in MicroAmp® 96-well Thin-Walled Reaction Plates (4346906, Life Technologies™). The plates were then centrifuged briefly to collect the contents at the bottom of the wells, before being placed into a Veriti® Thermal Cycler (4375786, Applied Biosystems®, Life Technologies™). The samples were incubated at 37^{o}C for 60 minutes, before stopping the reaction at 95^{o}C for 5 minutes and then holding at 4^{o}C. Once completed, the plates were stored at -20^{o}C.

II.IV.IV PCR Optimisation

Prior to analysing the samples, the qPCR was optimised using three different master mixes – TaqMan® Gene Expression Master Mix, Universal Master Mix and TaqMan® Environmental Master Mix 2.0. The positive control (sent by Nicolas Salez - positive control for qPCR (100 µl at 28CT)), was supposed to have a cycle threshold (CT) value of 28, but all three master mixes gave a CT value over 30. The TaqMan® Environmental Master Mix 2.0 gave the best CT value, and so was used for the qPCR.

II.IV.V qPCR

The primers and probe came in a preparation where they could not be immediately used and needed to be prepared. The probe had to be changed to a concentration of 2.5uM, and the

primers needed to be at a concentration of 5uM. These could then be stored and used in the qPCR.

The cDNA samples were then prepared in MicroAmp® Optical 96-Well Reaction Plate (N8010560, Applied Biosystems™, Life Technologies™). Specific fluC primers and probes for the M gene were used (Forward primer: CAT AAT TGA ACT TGT CAA TGG TTT TGT; Reverse primer: TTC AGG CAT AAT TGT GGT CTT TAT ATC T; probe: FAM-CTC GGC AGA TGG GAG AGA TGG TGT GTAMRA (28).) Each individual reaction was optimised for the primer and probe concentrations at a volume of 20.0µl (10µl of 2X TaqMan® Environmental Master Mix 2.0 (4396838, Applied Biosystems™, Life Technologies™), 2µl (500 nM) of each primer (Applied Biosystems™, Life Technologies™), 2µl (250 nM) of the probe (Applied Biosystems™, Life Technologies™) and 4µl DNA). For each plate positive and negative control ('no template controls') wells were also prepared, with the positive controls containing 4µl 'positive control', and the negative control containing 4µl of sterile PCR grade water (Sigma-Aldrich). The qPCR amplifications were then performed in a 7500 FAST Real-Time PCR system (Applied Biosystems®, Life Technologies™) with the thermo-cycling carried out at the following temperatures: one cycle of 95°C for 10 min and 45 cycles of 95°C for 15 s and 60°C for 1 min.

II.IV.VI Test for inhibition

An inhibition test was performed to check that there was nothing in the samples which would prevent or inhibit amplification. Five samples were chosen, which had appeared negative following qPCR. Each reaction was at a volume of 30µl, and contained: 15µl 2X TaqMan® Environmental Master Mix 2.0 (4396838, Applied Biosystems™, Life Technologies™), 3µl (500 nM) of both primers (Applied Biosystems™, Life Technologies™), 3µl (250 nM) probe

(Applied Biosystems™, Life Technologies™), 4µl Sample DNA (sterile PCR grade water (Sigma-Aldrich) used for controls), 2µl Positive Control DNA. The same process was then followed as for the qPCR.

II.V ELISA

II.V.I Total IgG ELISA optimisation

A direct ELISA was used with the intention of calculating the total IgG concentrations in the serum samples. In order to have reliable results, the ELISA protocol needed to be optimised. The method in section II.V.II was followed. Initially the sera were diluted at concentrations of 1:500 (2µl serum + 1ml carbonate buffer (CB)) and 1:1000, with the secondary antibody (ab6858, goat anti-human HRP, Abcam®) at a concentration of 1:2000 (1.0µg/ml in BB) and fish gelatine (2.5% (v/v) coldwater fish gelatine in 0.05% (v/v)) as the block. As these resulted in an excessively intense signal, the secondary antibody was tried at dilutions of 1:2000 and 1:4000. Again these samples were too strong and different blocks were compared

- Marvel (1% w/v Marvel milk powder in PBST (0.05%)), fish gelatine and bovine serum albumin (BSA) (1% w/v BSA in PBS). As the plates still developed too quickly the SureBlue™ (TMB Microwell Peroxidase Substrate solution (KPL)) used was compared to SureBlue Reserve™ (TMB Microwell Peroxidase Substrate (1-Component)).

Samples were then diluted to concentrations of 1:1000 and 1:4000; 1.10 000 and 1:20 000; 1:250 000, 1:500 000 and 1: 1 000 000; and finally $1:10^4$, $1:10^5$, $1:10^6$, $1:10^7$, $1:10^8$ and $1:10^9$ (over four different trials). Final results showed that the most successful combination was to use BSA as the blocking buffer, the secondary antibody at a concentration of 1:4000, and samples at dilutions of 1:100 000, 1:1 000 000 and 1:10 000 000.

II.V.II ELISA –Total IgG (Direct)

In order to be able to fully interpret the levels of fluC IgG antibodies, the total IgG antibody levels were also calculated.

Serum samples were diluted to three different concentrations in carbonate buffer (CB): 1:100 000, 1:1 000 000, 1: 10 000 000.

- 1µl + 1000µl CB = 1:1000 (A)
- 10µl A + 990µl CB = 1:100 000 (B)
- 100µl B + 900µl CB = 1:1 000 000 (C)
- 100µl C + 900µl CB = 1:10 000 000 (D)

Serum dilutions B, C and D were added to the plate in duplicate, allowing 10 samples to be analysed per plate.

Twenty four wells were filled with twelve doubling dilutions of IgG standards in duplicate. Firstly, to make up the IgG curve standards, 5µl IgG (1mg/ml) was added to 2ml carbonate buffer to give the first standard (2.5µg/ml). One ml carbonate buffer was then added to 11 more tubes, before adding 1ml of the previous IgG standard, creating 12 double dilution standards in carbonate buffer.

The remaining twelve wells only had carbonate buffer added, allowing for any background absorbance to be monitored.

100µl of each solution was added to each relevant well. The plate was then covered with Parafilm® and left to incubate overnight at 4^{o}C.

The following day the plate was sharply inverted to remove any excess unbound serum/IgG and blotted on blue paper to absorb any excess. The plate was then washed three times using PBST (0.05%) - 200µl/well. All wash stages were performed this way.

200µl blocking buffer (1% w/v BSA – 1g in 100ml PBS) was added to each well and left to incubate at 37^{o}C for one hour. The plate was then sharply inverted and washed as stated above.

Secondary antibody (Goat Anti-Human IgG (HRP), ab6858, Abcam®) was diluted to a concentration of 1:4000 in BB (2.5µl 2^{o}antibody in10ml BB) and 100µl was added to each well before incubating again at 37^{o}C for one hour. The plate was again sharply inverted and washed.

100µl SureBlue™ TMB Microwell Peroxidase Substrate solution (KPL) was added to each well and left to incubate at room temperature for four minutes. 100µl stop solution (1.6% v/v H_2SO_4 solution in water) was added to each well and then finally the absorbance was measured at 450nM using a Wallac Victor^{2}™ (Perkin Elmer) plate reader.

In order to estimate the concentration of IgG in each sample, absorbance values were plotted on the IgG standard curves on each plate. A line was drawn along the 'straight' aspect of the curve, with values at concentrations of 0.078125µg (7.11, -log (concentration of IgG)) and0.002441µg (8.61) used on each graph to allow direct comparison between plates. The average background absorbance was detracted from each sample in order to remove any non-specific contamination etc. Examples of the standard curves can be seen in Figures 12-15.

II.V.III ELISA – Influenza C antibodies (Indirect)

An ELISA was performed to detect antibodies in the patient's bloodstream, showing that they have previously been exposed to fluC. The method used was a protocol provided by Nicolas Salez, who has previously done work on fluC (28).

This method was also compared to a standard protocol used in our labs for detecting alpha haemolysin (AH) antibodies, but Salez's protocol appeared to only work for fluC antibodies.

Maxisorp® Plates (Nunc ImmunoMaxiSorp® F96, Thermo Scientific) were coated using the fluC antigen diluted to a concentration of 1:400 in PBS. 100µl was added to each of the test and control wells. For the standard curve, doubling dilutions of the human IgG were prepared, starting at 2.5µg/ml (5µl IgG (1mg/ml) added to 2ml carbonate buffer). Twelve wells were coated with 100µl of the IgG standards in carbonate buffer. The plate was then covered in Parafilm and wrapped in foil to avoid the light, then incubated overnight at 4^{o}C.

The following day, each plate was sharply inverted to remove any unbound antigen or IgG and then tapped on blue paper roll in order to absorb any excess liquid. Following this, the plates were washed three times using 200µl PBST (phosphate buffered saline + tween 0.05%) per well.

The next stage was to add the block; 200µl of the blocking buffer (1% Bovine Serum Albumin (BSA) in PBS) was added to each well, and then the plates were covered and incubated for two hours at 4^{o}C.

The plates were then washed again as outlined above. The serum samples were then prepared in a dilution of 1:400 in PBST + milk (2µl serum in 800µl PBST +milk 1%). One hundred µl of the diluted serum was then added to the relevant wells. Twelve wells (four wells of each serum) on each plate contained Sigma serum (H4522 Sigma**,** Human Serum from human male AB plasma, USA origin, sterile-filtered, Sigma-Aldrich®), positive control serum and negative control serum, also diluted to 1:400. At least six wells on each plate were used as 'background' wells to enable the plate to be checked for any non-specific binding or contamination. The remaining wells had 100µl of PBST + milk added. Again the plates were covered and incubated for 45-50 minutes at 37^{o}C.

Following the incubation, the plates were washed again. The secondary antibody (Goat Anti-Human IgG (HRP), ab6858, Abcam®) was prepared in PBST + milk to a dilution of 1:2000.

Per plate 10ml of PBST + milk was mixed with 5µl of the 2^{o} antibody. This mixture was then added to each well in quantities of 100µl. The plates were then covered and incubated for a final time at 37^{o}C for one hour.

The plates were washed for a final time, before 100µl of SureBlue™ TMB Microwell Peroxidase Substrate solution (KPL) was added to each well. The plates were then left for four minutes, before adding 100µl of stop solution (1.6% v/v H_2SO_4 solution in water) to each well. Finally, the plates were read at 450nm for one second using a Wallac Victor^{2}™ (Perkin Elmer) plate reader.

All absorbance values had the average 'background' absorbance subtracted, in order to get a more accurate impression of the true absorbance. These values were then plotted on the standard curve for the relevant plate, to allow estimations of the IgG concentrations. The same points of reference (0.078125µg (7.11) and 0.002441µg (8.61)) were used on each curve to allow for comparison between plates.

II.V.IV ELISA Sensitivity and specificity

When compared to a virus neutralisation test, the ELISA protocol had a sensitivity of 66%, a specificity of 86% and a positive predictive value of 91% (28). This shows that for those samples that are classed as positive, they have a high probability of actually having fluC antibodies; however despite the high specificity and positive predictive values, the ELISA actually underestimates the number of positive samples and therefore the seroprevalence rates.

II.V.V Analysis of positive and negative control sera

The ELISA protocol in section II.V.II 'ELISA –Total IgG (Direct)' was also applied to both the positive and negative control sera, which allowed for a greater analysis and understanding of the control sera.

An ELISA following the protocol in section II.V.III 'ELISA – Influenza C antibodies (Indirect)' was also carried out comparing three different elements. The new antigen was tested at dilutions of 1:400 and 1:800, in order to confirm that both antigen preparations sent by Nicolas Salez were at the same concentration. The positive and negative controls and the Sigma serum were also applied to wells without the antigen to check for any non-specific binding.

II.VI Agar Plates

Agar plates were prepared to detect the presence of coagulase-positive *S. aureus* on the swabs. Baird-Parker Agar Base was dissolved into distilled water at a ratio of 58g BP /950ml distilled water. The solution was then autoclaved (at 121°C, 15psi, for 15minutes) to sterilise it. A Bunsen burner was lit to create an updraft and to help prevent any bacteria from falling into the prepared plates. Following this, the solution was then cooled to 50^{o}C before adding the egg yolk (tellurite) emulsion, warmed to 50^{o}C, at a ratio of 12.5ml egg yolk emulsion/250ml Baird-Parker. Once the plates had set, they were stored upside-down in the cold room until required. Each plate was streaked with a nasal/nasopharyngeal swab from a different participant, and then incubated at 37^{o}C, again upside-down, for 24 to 48 hours. Any positive results appeared as grey/black in colour and had a clear halo around them.

II.VII Deep Sequencing

Any sample that appeared positive for fluC following PCR was prepared and sent to IBERS Phenomics Centre, Aberystwyth University, for deep sequencing. A selection of other samples were sent for deep sequencing to allow further analysis of the different types and groups of participants. Double-stranded cDNA was required for analysis, and this was prepared from the raw RNA extracted from the samples using the 'SuperScript® Double-Stranded cDNA Synthesis Kit' (11917-010 Invitrogen™, Life Technologies™). For those

samples which were to be pooled for deep sequencing, the RNA in the samples was mixed prior to conversion to cDNA.

Nine different Eppendorf tubes, labelled A – I, were sent for deep sequencing, with each Eppendorf tube containing a different variety of samples (see appendix 1).

- A - Paediatric and symptomatic (lower RNA reading on NanoDrop)
- B - Paediatric and symptomatic (higher RNA reading on NanoDrop)
- C - Adult, high fluC IgG, symptomatic
- D - Adult and fever
- E - Asymptomatic, low total IgG
- F - Asthmatics (higher fluC IgG)
- G - COPD
- H - Positive PCR
- I - Positive PCR

Unfortunately due to delays, the results from the deep sequencing are yet to be returned and analysed.

II.VIII Bioinformatics

II.VIII.I Software

Various bioinformatics software was used in order to analyse all of the different samples.

- NCBI Influenza Virus Resource (92)
    - Basic Local Alignment Search Tool (BLAST®) (93)
- Artemis (Sanger Institute) (94)
- Path-O-Gen v1.4 (95, 96)
- Molecular Evolutionary Genetics Analysis (MEGA) 6.06 (97)

- Bayesian Evolutionary Analysis Utility (BEAUti) v1.8.0 (95)
- Bayesian Evolutionary Analysis Sampling Trees (BEAST) v1.8.0 (95)
- Tracer v1.6 (98)
- TreeAnnotator v1.8.0 (95)
- FigTree v.1.4.2 (99)

The NCBI National Center for Biotechnology Information (NCBI) PubMed database (100) was also used as the basis for the literature review.

II.VIII.II Phylogenetic trees

Using the database GenBank (NIH genetic sequence database) (101) every fluC strain for each segment of the virus was downloaded (in FASTA format). These data were then able to be analysed using the bioinformatics software mentioned previously.

Initially, every segment of the fluC virus was visualised using the Artemis software (94).

The data were aligned by ClustaIW using MEGA (97), and any strains found in hosts other than humans were removed. The best model was then determined using MEGA (97), before a 'best fitting tree' was created.

In addition, the best fitting clock, R_0 values and Bayesian skyline (102) were all analysed using the bioinformatics software mentioned previously. Path-O-Gen was used to check for clock-like behaviour and to derive the approximate time of the common ancestor (96).

III Results

III.I Recruitment summary

Overall, 148 participants were recruited in the study. Figures 8-10 show the various groups present. 143 participants provided a nasal or nasopharyngeal swab (25 nasal, 118 nasopharyngeal) and 129 participants gave a serum sample (101 venous, 28 capillary).

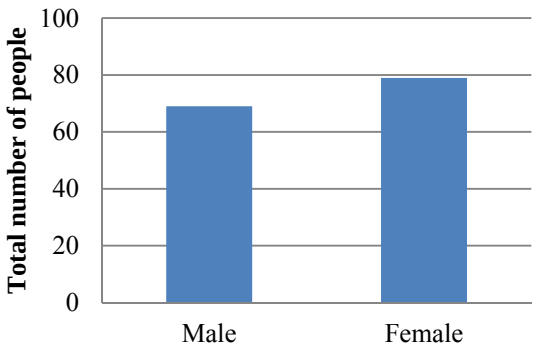


Figure 8 – Gender of participants recruited

69 males and 79 females were recruited into the study.

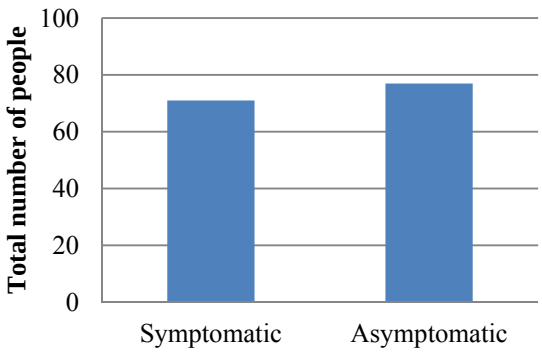


Figure 9 – Symptomatic and asymptomatic participants recruited

71 symptomatic and 77 asymptomatic participants were recruited for this study.

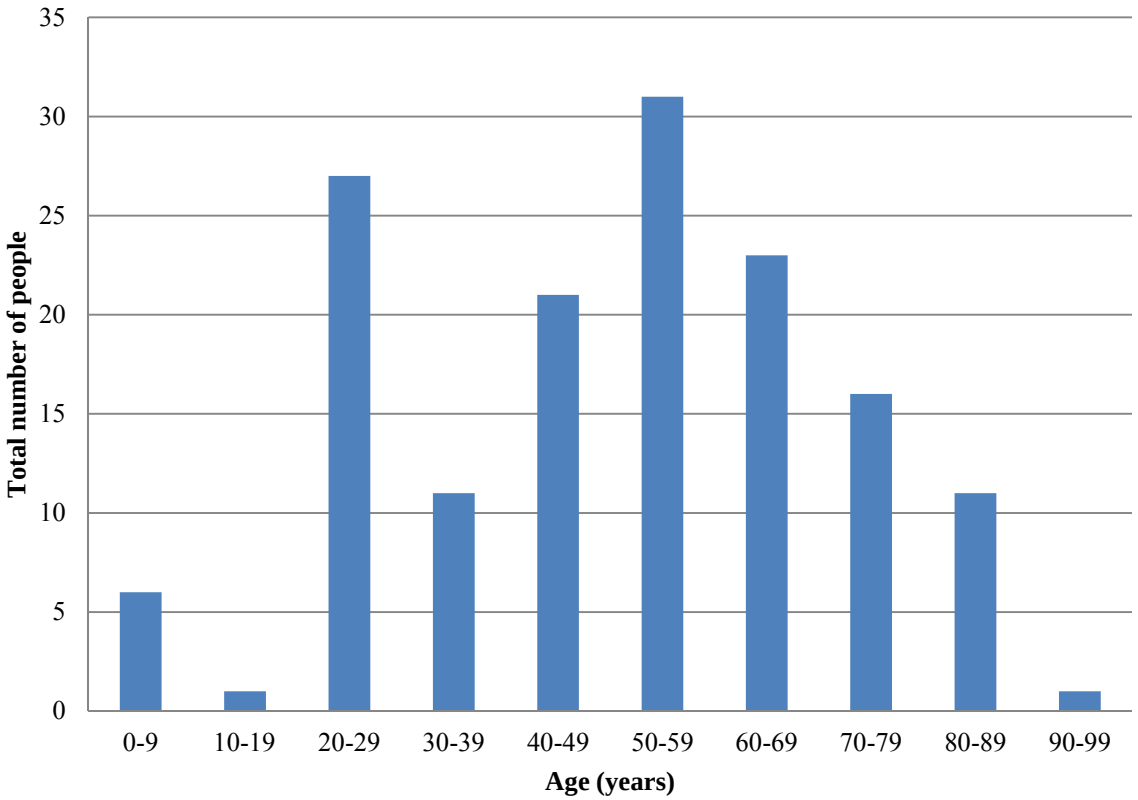


Figure 10 – Age of participants recruited

More participants were recruited in the 20-29 and 50-59 age groups. Participants (n) were binned into 10-year spans of 0-9 (n=6), 10-19 (n=1), 20-29 (n=27), 30-39 (n=11), 40-49 (n=21), 50-59 (n=31), 60-69 (n=23), 70-79 (n=16), 80-89 (n=11) and 90-99 (n=1).

Table 4 - Symptoms experienced by participants

The most common symptoms described by 'symptomatic' participants:

Symptom	No. of participants
Cough	51
(Productive)	23
Coryza	42
Dyspnoea	26
Sore throat	24
Malaise/fatigue	19
Fever/Rigors	15
Sinusitis	13
Headache	12
Myalgia/weakness	11
Chest pain (pleuritic)	11
Nausea/vomiting	6
Decreased appetite	4
Wheeze	3
Diarrhoea	3
Drowsiness/confusion	3
Otalgia	2
Haemoptysis	2
Tachypnoea	1
Abdominal pain	1
Dysphagia	1
Dizziness	1
Haematuria	1
Hypotension	1

III.II PCR results

Two participants out of 148 (1.35%) were detected as positive for fluC via PCR. Those participants were K081 and K087, as shown in Figure 11. The samples contain the fluC cDNA and therefore, as each cycle completes, the genetic material increases, which is detected by the PCR machine, therefore giving off a signal. These two positive samples were then sent off for further deep sequencing.

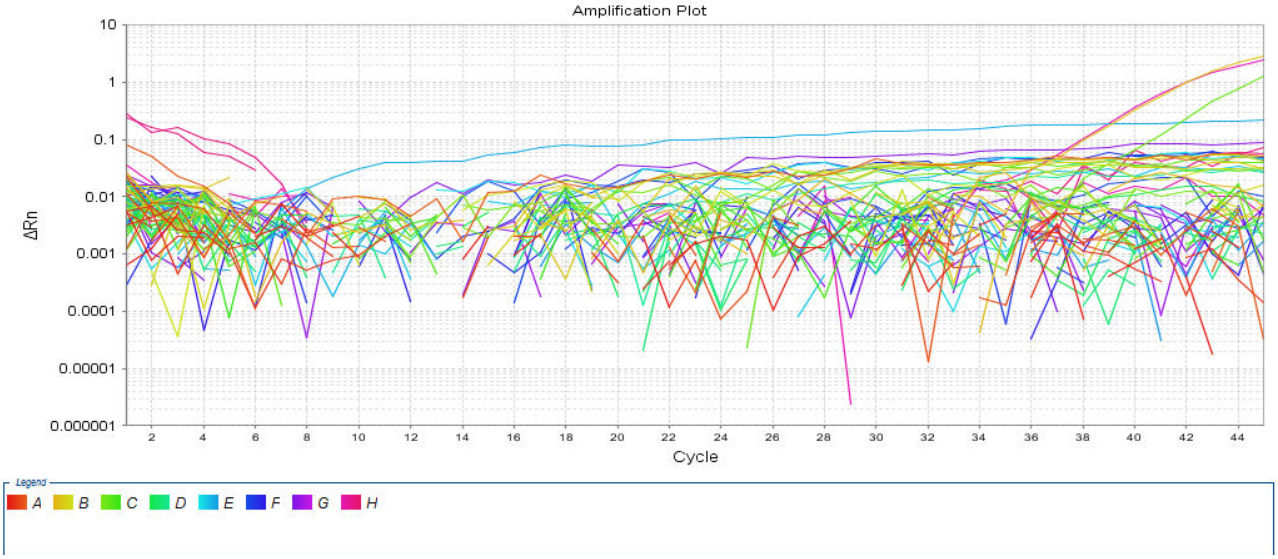


Figure 11 - PCR Amplification plot, plate 2

The positive samples K081 and K087 and the positive control sample are shown. Each well containing a sample gives off a signal, shown in the Figure as a coloured line: row A - red line, row B - yellow line, row C - light green line, row D - dark green line, row E - light blue line, row F - dark blue line, row G - purple line, row H - pink line, as shown in the legend above. As each cycle completes, any samples containing fluC will increasingly contain more cDNA and therefore will give off a stronger signal. The positive control (pink/legend H) and sample K081 (yellow/legend B) both begin to increase from a cycle threshold (CT) value of around 34, and overlap slightly, whereas sample K087 (green/legend C) begins to increase nearer to 40 CT, and then all three exponentially increase for the remainder of the cycles.

III.III ELISA results

As mentioned in section II.V.I 'Total IgG ELISA optimisation', several ELISAs were attempted in order to optimise the protocol for detection of total IgG antibodies, prior to testing the serum samples. Several different methods were trialled prior to finding a successful protocol. Initially the serum samples were tested at dilutions of 1:500 and 1:1000, with the secondary antibody diluted to 1:2000 and fish gelatine used as the block. These samples were far too strong and so the next method tried involved trialling the secondary antibody at two concentrations –1:2000, as before, and 1:4000. The serum samples were again left at concentrations of 1:500 and 1:1000 in order to keep all other elements the same. Again, all samples were still too strong and so, keeping all other constants the same, different blocks were tested; 'Marvel' milk powder, fish gelatine and BSA. Once again the samples developed too quickly, but appeared to be slightly more effective with BSA. It was then noted that SureBlue Reserve had been used to develop the plates, which is 50% more sensitive than SureBlue, and so the next plates tested compared SureBlue and SureBlue Reserve. Despite this, there wasn't much difference between the plates and so the next step was to try diluting the samples further. The samples were diluted to 1:1000 and 1:4000 in the first trial and 1.10 000 and 1:20 000 in the second, all of which were still too strong. A paper was then found which tested total IgG levels in rat sera at a dilution of 1:200 000 (103) and so it became apparent that the samples were still at least 10 times too concentrated. The samples were then trialled at dilutions of 1:250 000, 1:500 000 and 1: 1 000 000 which gave better results, but most samples were still too strong. A specific ELISA for AH IgG antibodies was trialled at dilutions of 1:10, 1:100, 1:1000, 1:10 000 1:100 000 and 1:1 000 000. Some AH IgG could still be detected at the weakest dilution, suggesting that the previous total IgG dilutions were still too strong. Finally the samples were tested at six different dilutions in order to get a full picture of what was going on. The dilutions used

were: 1:10 000, 1:100 000, 1:1 000 000, 1:10 000 000, 1:100 000 000 and 1:1 000 000 000. Samples around 1:1 000 000 fell nicely on the straight part of the curve, and so it seemed to be the best dilution to use. It was then decided to test all serum samples at dilutions of 1:100 000, 1:1 000 000 and 1:10 000 000 in order to account for any samples which may have more or less IgG than expected. The block chosen was BSA and the secondary antibody was used at a concentration of 1:4000. A few mock ELISAs were also trialled in order to make sure that the standard curves were accurate, so that the IgG concentrations could be estimated in the samples. Once this process was complete, all of the serum samples were then tested so that a value could be calculated for the total IgG concentration of each participant. Optimising the protocol also allowed for perfecting pipetting skills etc., which can be demonstrated using concentration curves. An indirect ELISA protocol was also trialled testing for AH antibodies, as a practice for the fluC ELISA. Figures 12-15 show standard concentration curves taken from various ELISA plates. Figures 12 and 13 represent the IgG curves from a plates looking at total IgG concentrations, Figure 14 pictures an IgG curve from an AH IgG plate and Figure 15 shows a standard IgG curve from a fluC IgG plate.

The curves each demonstrate doubling dilutions of known IgG concentrations, which when plotted on a graph shows the typical ‘sigmoid’ shape. These curves can then be used to estimate the levels of IgG in the samples.

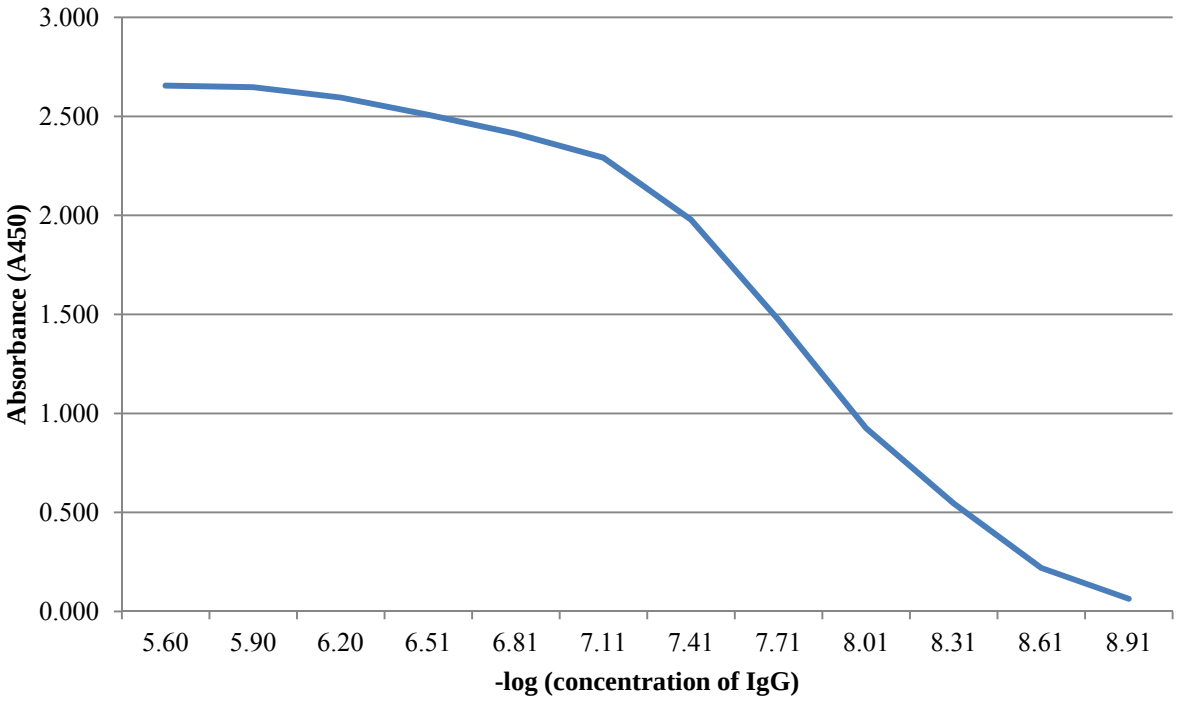


Figure 12 – Standard IgG Concentration Curve (26/3/15) - Total IgG Plate

The standard IgG curve for the ELISA plate comparing serum sample total IgG concentrations of 1:250 000, 1:500 000 and 1:1 000 000

Table 5 - ELISA results establishing total IgG concentrations (10/4/15)

The results from the total IgG ELISA comparing serum samples at dilutions of 1:10 000, 1:100 000, 1:1 000 000, 1:10 000 000, 1:100 000 000, and 1:1 000 000 000. These results were also repeated and confirmed. Red indicates high levels of IgG in a sample, and green indicates a low level of IgG. Each coloured cell indicates a well on the ELISA plate. The two rows of 'curve' wells contained samples with known levels of IgG, in doubling dilutions, which when plotted produces a standard IgG curve of known concentrations which allows for estimation of the IgG in each of the serum sample wells.

	(1:10 000)		(1:100 000)		(1:1 000 000)		(1:10 000 000)		(1:100 000 000)		(1:1 000 000 000)	
Sigma*	2.767	2.702	2.707	2.698	2.227	2.170	0.713	0.792	0.194	0.173	0.860	0.728
FBS*	0.077	0.075	0.079	0.071	0.080	0.083	0.074	0.061	0.069	0.077	0.068	0.063
Serum*	2.795	2.733	2.645	2.720	1.716	2.068	0.572	0.548	0.200	0.212	0.176	0.195
	2.735	2.740	2.481	2.399	1.078	1.047	0.384	0.268	0.132	0.161	0.116	0.125
	2.811	2.797	2.755	2.698	2.432	2.492	0.841	0.886	0.986	1.126	0.796	0.772
Background	0.130	0.105	0.089	0.079	0.074	0.074	0.070	0.070	0.071	0.072	0.066	0.430
Curve	2.811	2.869	2.790	2.764	2.715	2.748	2.711	2.747	2.690	2.580	2.495	2.423
Curve	2.171	2.145	1.768	1.716	1.159	1.241	0.783	0.763	0.452	0.465	0.286	0.289

*dilutions 1:10 000 - 1:1 000 000 000 (in duplicate), left to right

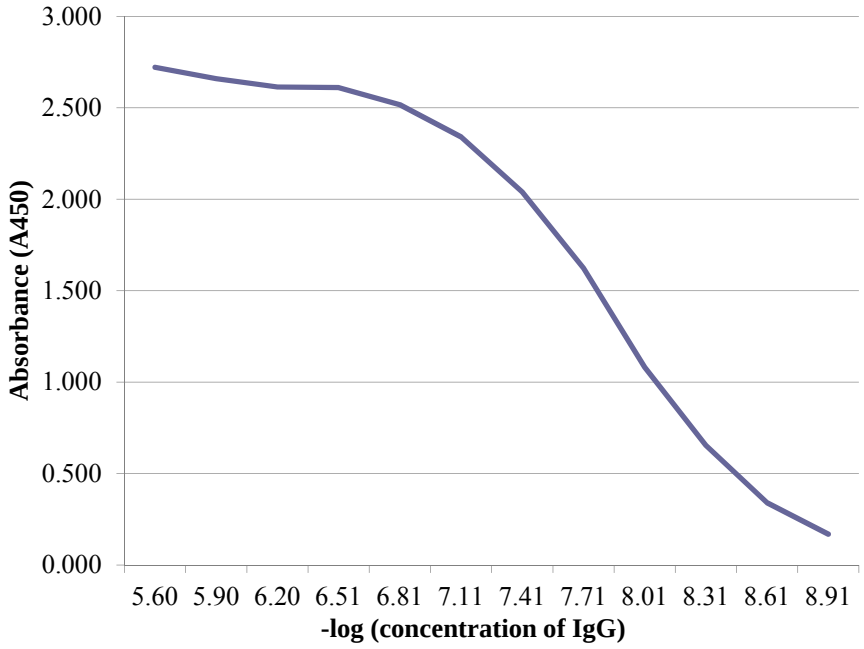


Figure 13 - Standard IgG Concentration Curve (10/4/15) - Total IgG Plate

This shows the standard IgG curve for the ELISA plate shown in Table 5.

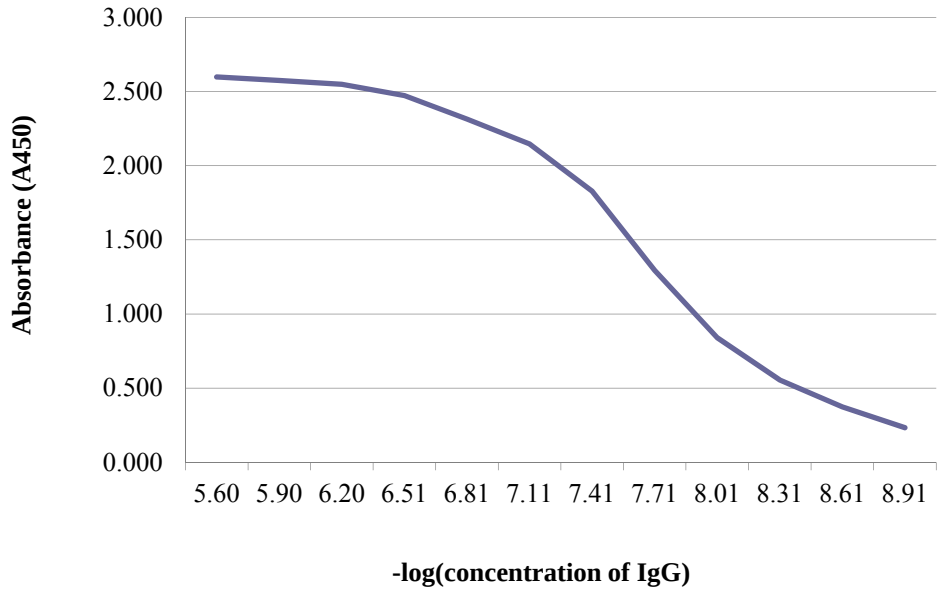


Figure 14 - Standard IgG Concentration Curve (21/4/15) - AH IgG Plate

This is the standard IgG curve for one of the AH ELISA plates, looking for specific IgG antibodies. Every venous serum sample was given an AH IgG antibody value.

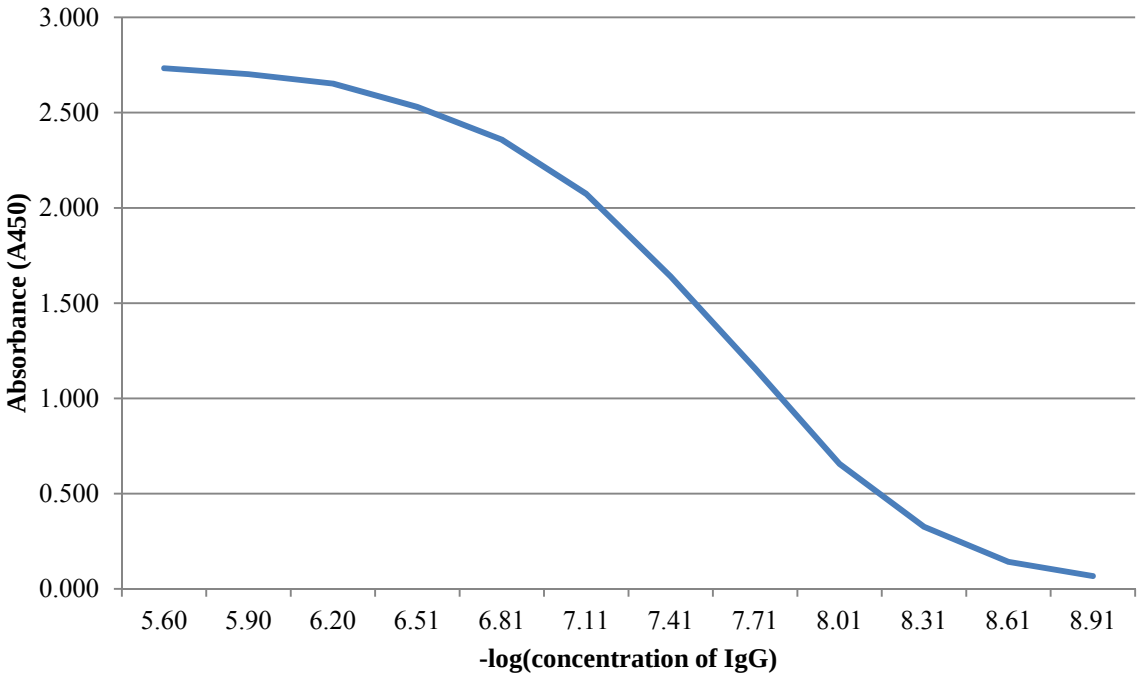


Figure 15- Standard IgG Concentration Curve (21/4/15) - FluC IgG Plate
This shows the standard IgG curve from a fluC IgG-specific ELISA plate.

III.III.I Influenza C IgG antibody levels

Table 6 (see appendix 2) shows the fluC IgG antibody results. Overall **106 serum samples were seropositive**, and **23 samples were seronegative**, giving a seropositivity of **82%**. Estimates were based upon the standard IgG concentration curves for each plate, with the maximum and minimum values (0.078125µg (7.11) and 0.002441µg (8.61)) remaining the same for each plate. Any sample which had an absorbance value above the maximum point on the line had to be given as an estimate above a set value. The calculations used to determine whether a sample was positive or negative are shown in section 'III.III.IV Statistics' and 'Figure 24'. A value of '-' indicates that no sample was provided.

Table 6 - Influenza C IgG Concentration (see appendix 2)

Figures 16, 18 and 19 show the mean fluC IgG antibodies in various population groups. The age distribution of fluC antibodies is shown in Figure 16, whereas Figure 18 compares the distribution of fluC IgG antibodies between different genders, and Figure 19 shows the distribution of IgG antibodies between the symptomatic and asymptomatic groups. There was little difference between the male and female participant IgG levels; however the symptomatic group had slightly higher levels of fluC IgG antibodies than the asymptomatic. Overall the variety of fluC IgG levels between the different age groups was small, with most age groups having similar levels of antibodies. The 10-19 years age group had lower levels of antibodies, but as there was only one participant in this group it is difficult to comment on. Figure 17 shows the age distribution of the mean fluC IgG antibodies in both the symptomatic and asymptomatic groups.

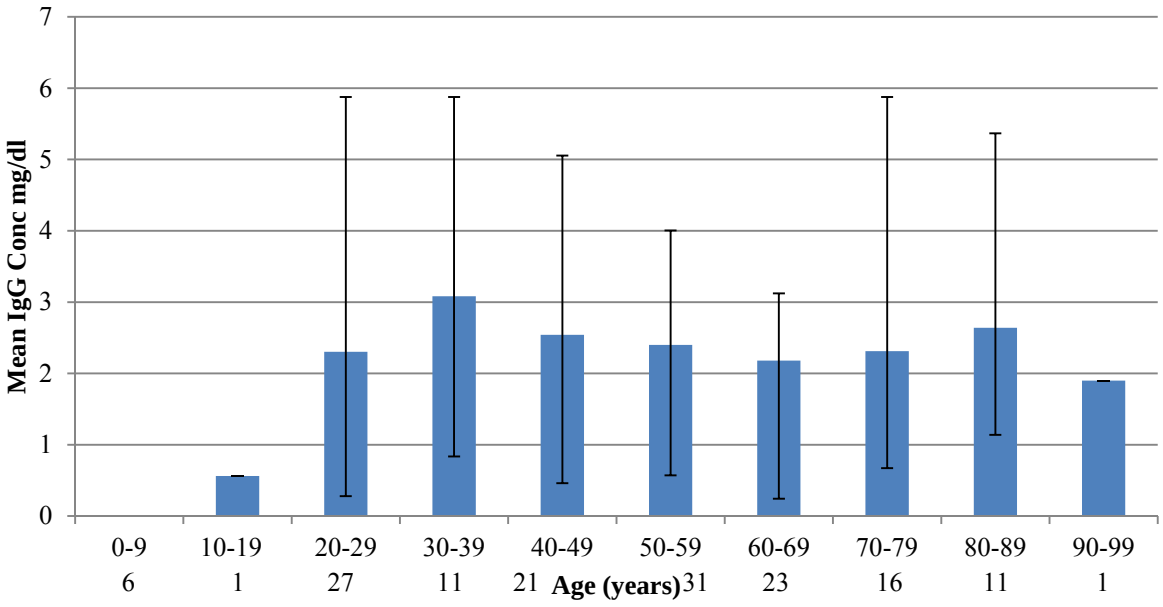


Figure 16- Age distribution of fluC IgG antibodies

The number below the age group on the x axis indicates the number of participants in that group. No participants in the 0-9 years group provided a serum sample. The bars indicate the variance in the data for the age group.

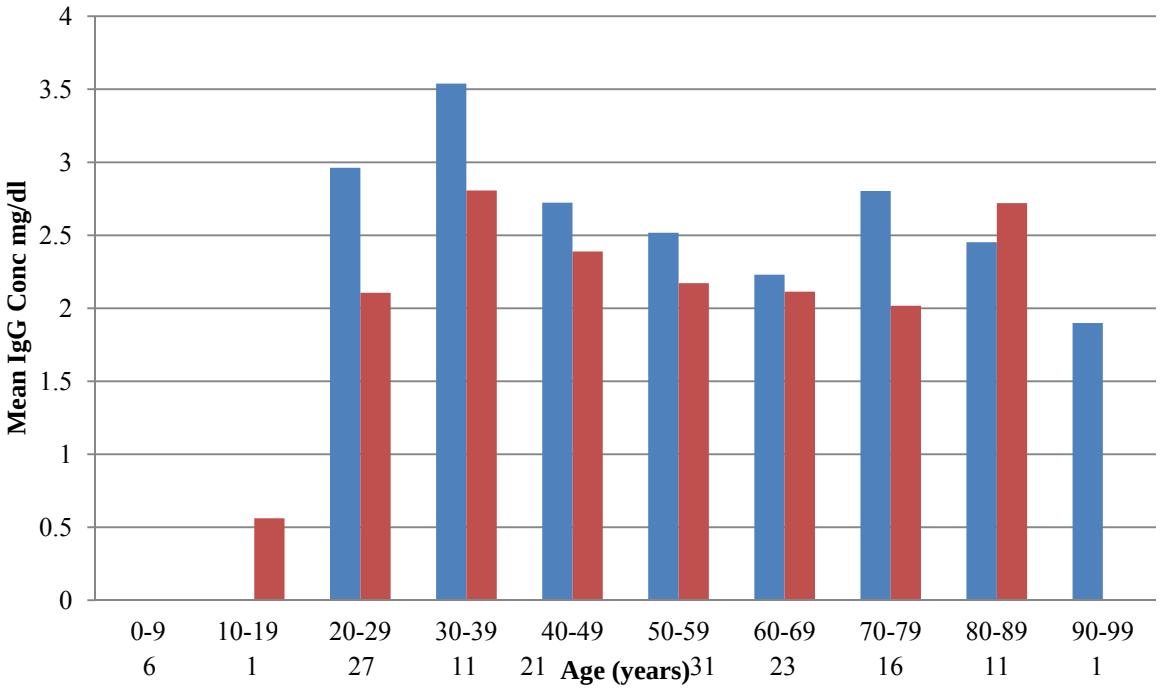


Figure 17 - Age distribution of fluC IgG antibodies - symptomatic vs. asymptomatic
The red bars indicate the asymptomatic participants and the blue bars indicate the symptomatic participants. No participants in the 0-9 age group provided a serum sample. The number below each age group on the x axis indicates the number of participants in that group.

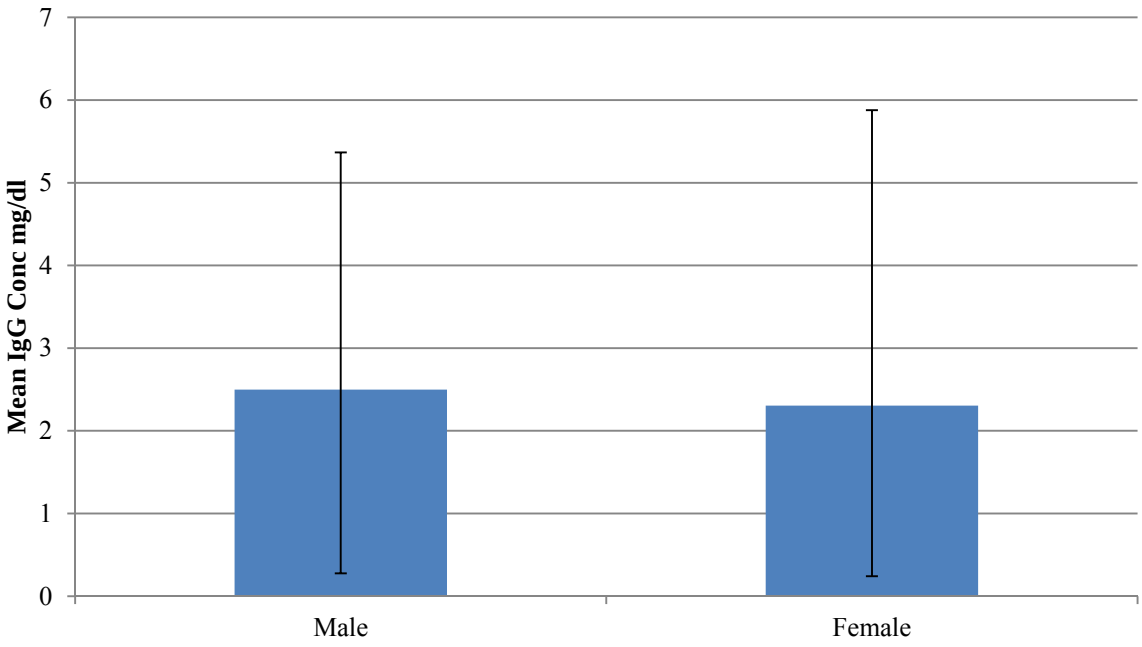


Figure 18 - Gender distribution of fluC IgG antibodies

The mean IgG concentration for the male participants was 2.50mg.dl and the mean IgG concentration for the female participants was 2.31mg/dl, showing that there is not a great difference between the genders. The error bars indicate the variance in the data.

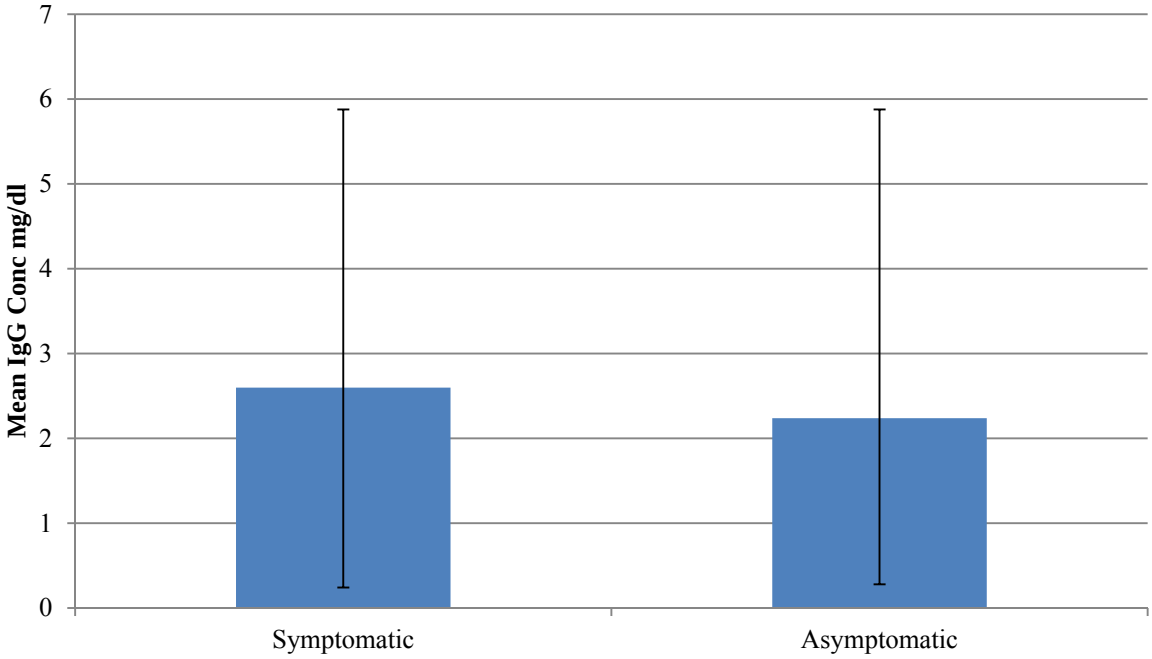


Figure 19 - Symptomatic and asymptomatic fluC IgG antibodies

The mean IgG concentration for the symptomatic participants was 2.60mg.dl and the mean IgG concentration for the asymptomatic participants was 2.24mg/dl, which shows that despite the symptomatic participants have higher levels of fluC IgG antibodies, it is not greatly different from the asymptomatic participants. The error bars demonstrate the range of the data.

III.III.II Antigen, positive and negative control serum tests

Some additional tests were done on the antigen and positive and negative controls sent by Nicolas Salez (Aix-Marseille Université), in order to confirm exactly what the samples contained.

The total IgG values for both serum samples were checked, to confirm that the negative serum was a true negative and not merely that the person who provided the sample was immunocompromised or had another similar problem. The values detected are in Table 7, which can be compared to those of the serum samples in Table 10.

Table 7 - Total IgG Concentration in positive and negative controls

This table shows the total IgG concentrations in mg/dl for the positive and negative control sera provided by Nicolas Salez.

	Total IgG concentration (mg/dl)
Positive control	1822.632
Negative control	2047.266

Two different antigen preparations were used for detecting the fluC antibodies. Both were sent by Nicolas Salez, but at different times, and so needed to be compared in order to prove that the samples had been analysed in the same manner. Tables 8 and 9 show the absorbance

values for the antigens when diluted to a concentration of 1:400, and the protocol outlined in section II.V.III 'ELISA – Influenza C antibodies (Indirect)' was followed.

Table 8 - Old antigen absorbance (A_{450}) values

This table shows the mean absorbance (A_{450}) values for the Sigma serum and positive and negative controls, when tested using the old antigen. The standard deviation (SD) and standard error (SE) are also shown.

Old antigen			
	Sigma	Positive	Negative
Mean	2.214	2.591	0.836
SD	0.103	0.140	0.144
SE	0.023	0.031	0.032

Table 9 - New antigen absorbance (A_{450}) values

This table shows the mean absorbance (A_{450}) values for the Sigma serum and positive and negative controls, when tested using the new antigen. The standard deviation (SD) and standard error (SE) are also shown.

New antigen			
	Sigma	Positive	Negative
Mean	2.206	2.591	0.996
SD	0.100	0.100	0.189
SE	0.038	0.038	0.071

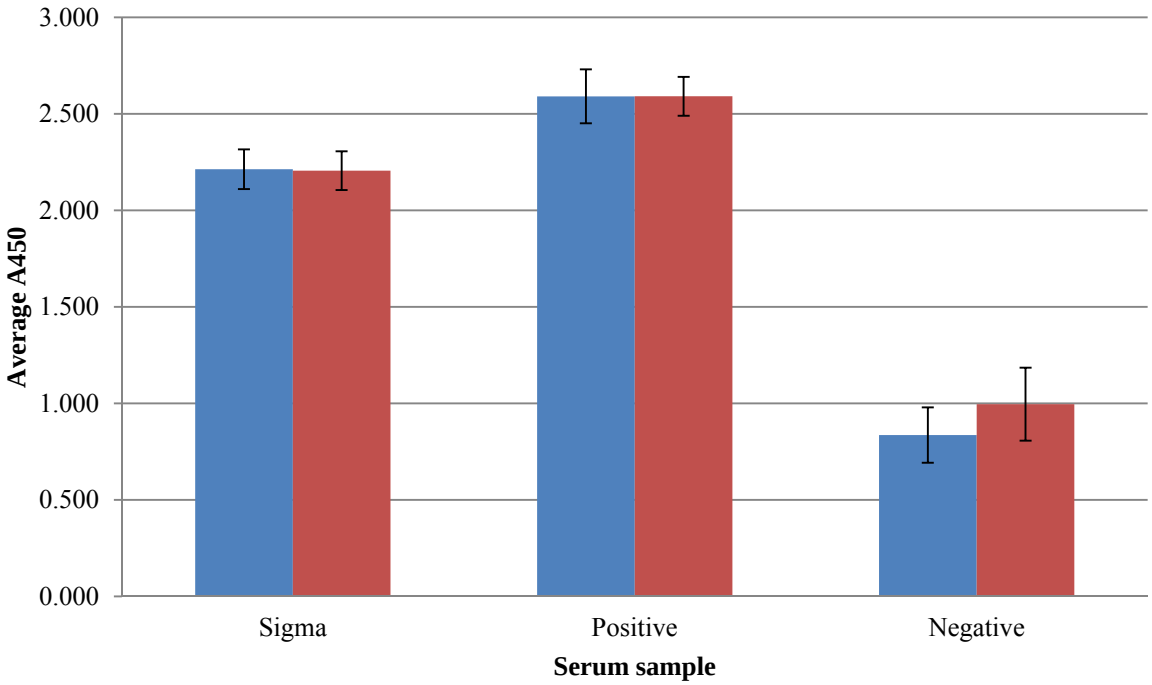


Figure 20 - Comparison of absorbance values in old and new antigen

Comparison of the absorbance values for the two different antigen preparations. The error bars show one standard deviation either side of the mean. The blue lines represent the old antigen and the red lines indicated the new antigen. The error bars show one standard deviation above and below the mean.

A T test was also done to confirm whether the difference between the antigens was significant. When comparing the positive control sera and the Sigma serum (H4522 Sigma**,** Human Serum from human male AB plasma, USA origin, sterile-filtered, Sigma-Aldrich®), between the new and old antigens, the results were not significant at $p<0.05$. The negative control sera were, however, significantly different at $p<0.05$, but were not significant at $p<0.01$.

III.III.III Total IgG antibody results

Table 10 contains the total IgG antibody results for all of the serum samples. Again, estimates for the concentrations were calculated using the IgG standard concentration curves for each

plate (points at 0.078125µg (7.11) and 0.002441µg (8.61) used). Generally, values from the serum dilutions at 1:1 000 000 were used. Any samples which had particularly high absorbance values were plotted on the curve using the 1:10 000 000 dilution instead, to allow for more accurate calculations. Any participants that did not provide a serum sample were given a value of '-' (these values were not included in any averages or statistics). The normal range for total IgG values were 639–1,349 mg/dl (54). Any values above this range were coloured red, and those below were coloured green. Overall, 71 samples were coloured red (above normal range), 22 samples were coloured green (below normal range), and the remaining 36 samples were coloured black i.e. within normal range.

Table 10 - Total IgG antibody results (see appendix 3)

Figures 21-23 show the distribution of the mean total IgG antibody levels in different population groups. Figure 21 shows the age distribution of total IgG antibodies, and it appears that most age groups have similar antibody levels. Although the 10-19 age group has a particularly low level and the 90-99 age group has a particularly high level, as only one participant was in each of these age groups these results cannot be relied upon. Figure 22 compares the symptomatic and asymptomatic group total IgG levels and Figure 23 compares the total IgG levels between male and females, however all of these groups give very similar results with little difference.

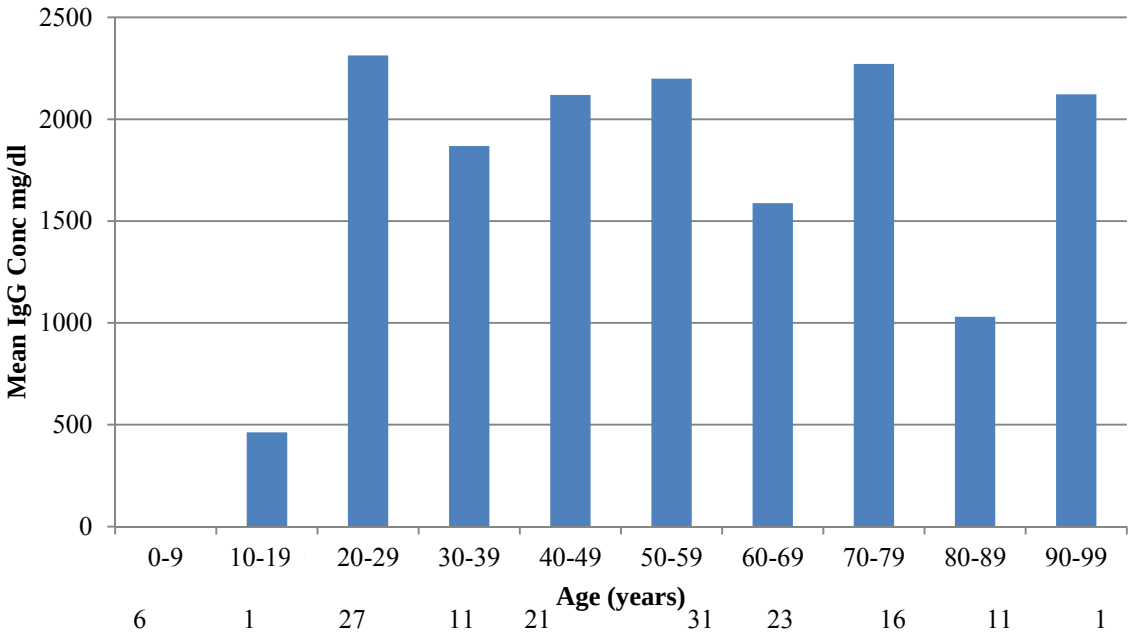


Figure 21 – Age distribution of total IgG antibodies

No participants in the 0-9 years group provided a serum sample. The numbers on the x axis below the age group show the number of participants in each age group.

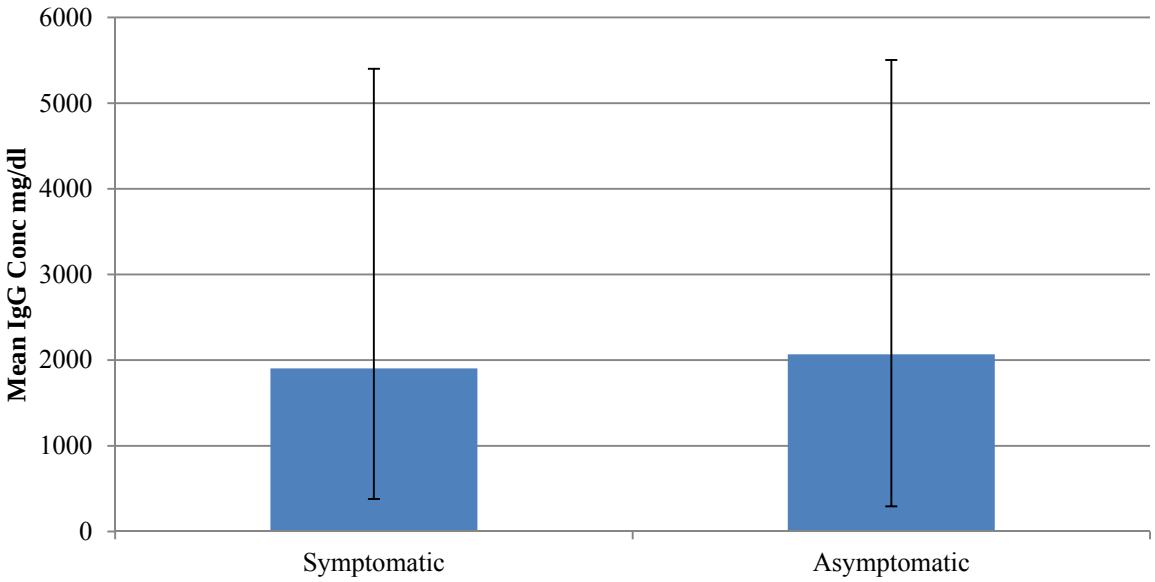


Figure 22 – Comparison of total IgG antibodies in symptomatic and asymptomatic participants

The mean total IgG level for the symptomatic participants was 1902.70mg/dl and for the asymptomatic participants was 2067.45mg/dl. The error bars show the variance in the data.

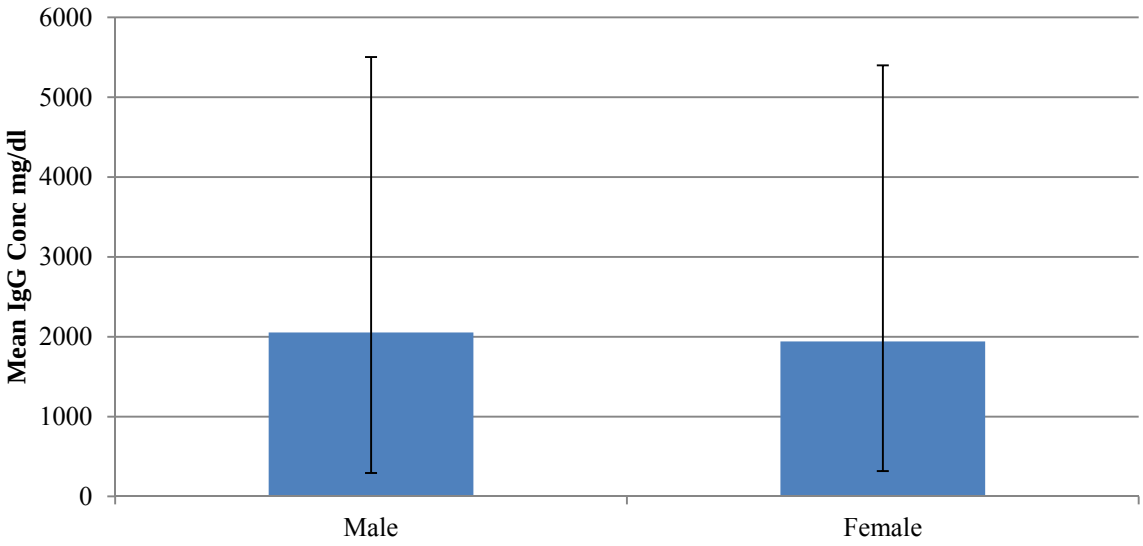


Figure 23 - Gender distribution of total IgG antibodies

The mean total IgG level for the male participants was 2054.13mg/dl and for the female participants was 1941.30mg/dl. The error bars demonstrate the variance in the data.

Table 11 - Comparison of FluC and Total IgG Antibodies (see appendix 4)

Table 11 compares the fluC and total IgG antibodies for each serum sample.

III.III.IV Statistics

In order to determine the cut-off values for positive and negative fluC antibody results, the values had to be analysed. As the antigen and positive and negative control sera were the same as used in the paper by Salez *et al.*(28) the data from this study were able to be directly compared with that in France. Every one of the French plates determined whether a sample was positive or negative using the same formula. All of the negative sera were averaged and then 2 standard deviations (SD) were added to the negative sera mean, with any samples that were above this value being classed as positive. In order to do so the plates needed to be 'normalised' to the Marseille plate, which was used as the standard for Salez's work. To do this, the following equation was used:

Value = A + (a – 'OD raw') x N/(b-a),

where A= 0.1037, b=Positive control, a= Negative control, N= 1.1011, and 'OD raw' is the raw optical density (OD)/absorbance value. Certain values remained constant between the different plates (A and N), whereas others varied to allow for the different conditions of the plates (a, OD raw and b). Any sample with a value >0.733 was classed as positive (In Salez's paper, a value >0.797 was classed as positive (28)). As there was the same standard positive value for all plates, an OD value may give a positive sample on one plate and a negative on another, due to the different values used in the formula. Figure 24 shows the positive and negative samples using this equation and plotting these values against age.

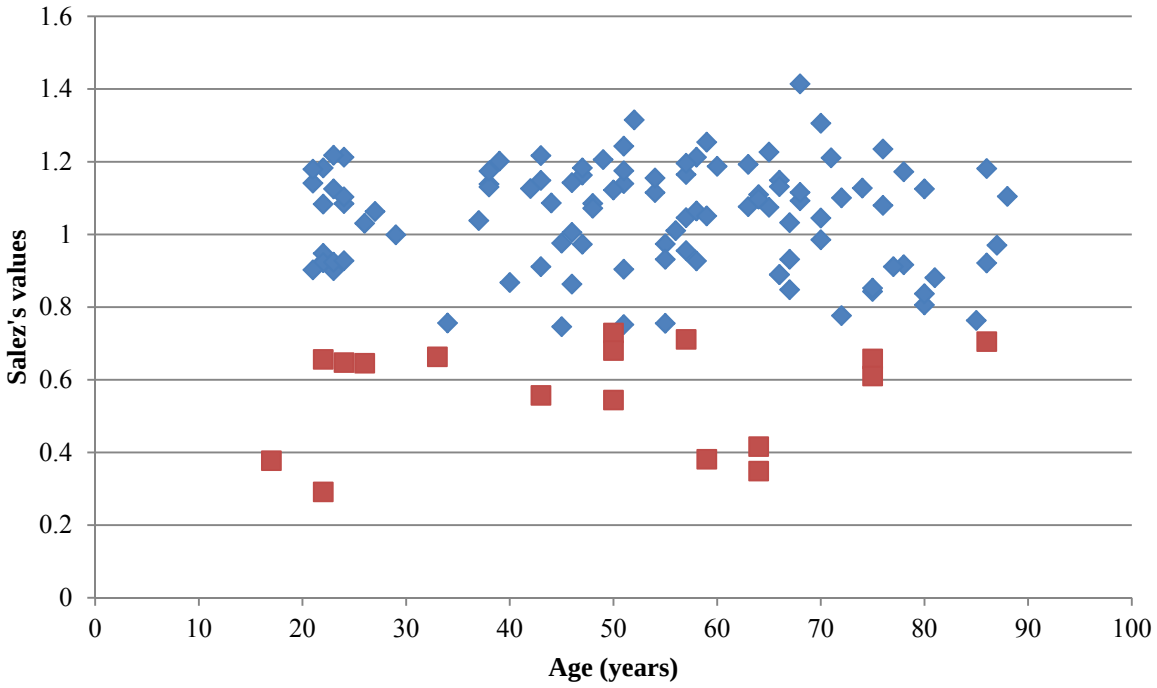


Figure 24 – Age distribution of positive and negative serum samples, as calculated by Salez's formula

Blue indicates a positive sample (>0.733), of which there were 106 samples, and red indicates a negative sample (<0.733), of which there were 23 samples.

An alternative method was also trialled to calculate positive and negative samples. It is defined that 95% of the data falls within 1.96 SD of the mean. This basis was used to decide whether or not a serum sample was positive or negative. Any sample which fell within 1.96

SD of the positive control mean was classed as positive, and any sample which fell within 1.96SD of the negative control mean was classed as negative, and all the others were classed as 'in-between'. This would give a positive control mean of 4.46 mg/dl (1.45-7.47mg/dl). Figure 25 and Table 12 show the mean fluC IgG concentrations for the positive and negative controls and also the Sigma serum. Error bars show one standard deviation either side of the mean. Figure 26 shows how the samples would be classified if this method was used.

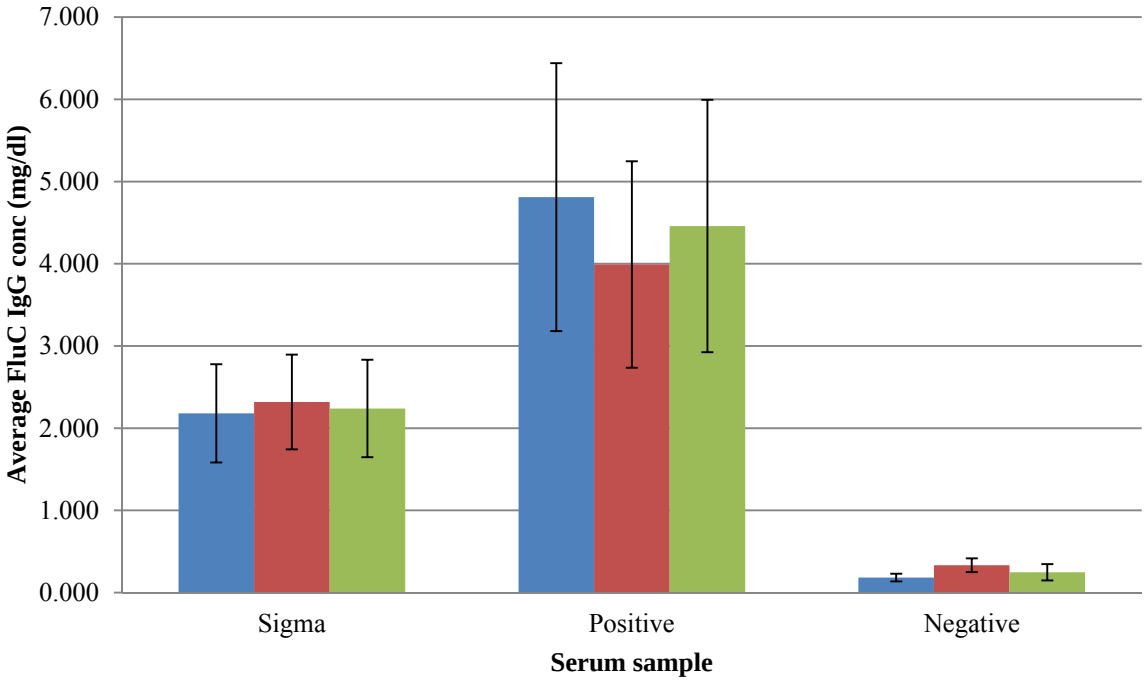


Figure 25 - Positive and negative control serum - FluC IgG concentrations

The blue lines indicate the old antigen, the red lines indicate the new antigen and the green line represents the mean of the old and new antigens. The error bars show one standard deviation either side of the mean.

Table 12 - Positive and negative control serum - FluC IgG concentrations

The mean fluC IgG concentrations (mg/dl) were calculated for the positive and negative sera and the Sigma serum. The standard deviation (SD) and standard error (SE) were also

calculated, and 1.96SD was both added and subtracted to the mean for each serum, in order to get an estimation of where 95% of the data should lie.

Combined antigen			
	Sigma	Positive	Negative
Mean	2.240	4.459	0.247
SD	0.592	1.536	0.099
SE	0.224	0.580	0.037
Mean +1.96SD	3.401	7.469	0.440
Mean - 1.96SD	1.079	1.449	0.053

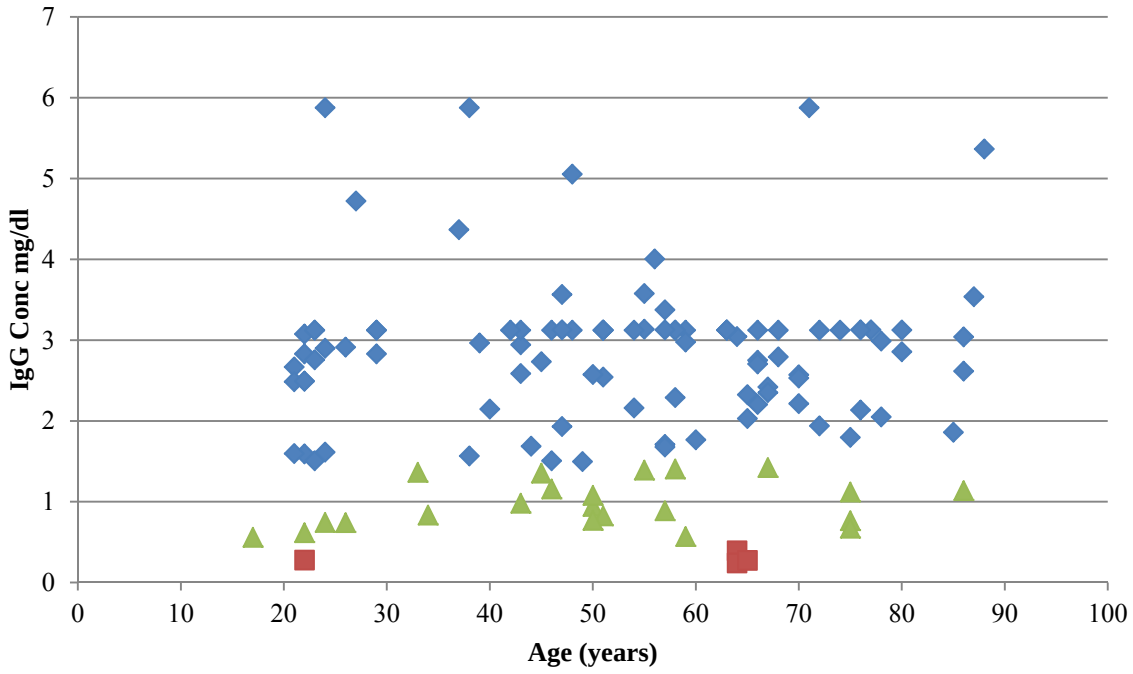


Figure 26 - Age distribution of positive and negative serum samples, as calculated by using 1.96 SD

This figure represents the data shown in Table 12. The positive samples (>1.449) are coloured blue, the negative samples (<0.440) are coloured red and the remaining samples in between the positive and negative regions are coloured green. This shows that 99 samples are positive, 4 are negative and the remaining 26 samples are in-between.

Table 13 - Chi square test - Male and Female, Symptomatic and Asymptomatic

This table shows the total number of male and female symptomatic and asymptomatic participants. A chi square test was then performed on this data.

	Symptomatic	Asymptomatic	Total
Male	25	36	61
Female	32	36	68
Total	57	72	129

The x^2 value for the chi square test was 0.481221, which shows that the probability of a large value of x^2 falls between 0.50 and 0.25. This means that the data can be considered to be randomised with respect to gender and clinical status.

In order to further analyse the data for any statistically significant differences, either T tests or Mann Whitney U tests needed to be used, based upon whether the data are normally distributed or not. The positive fluC IgG results were plotted against age, as shown in Figure 27.

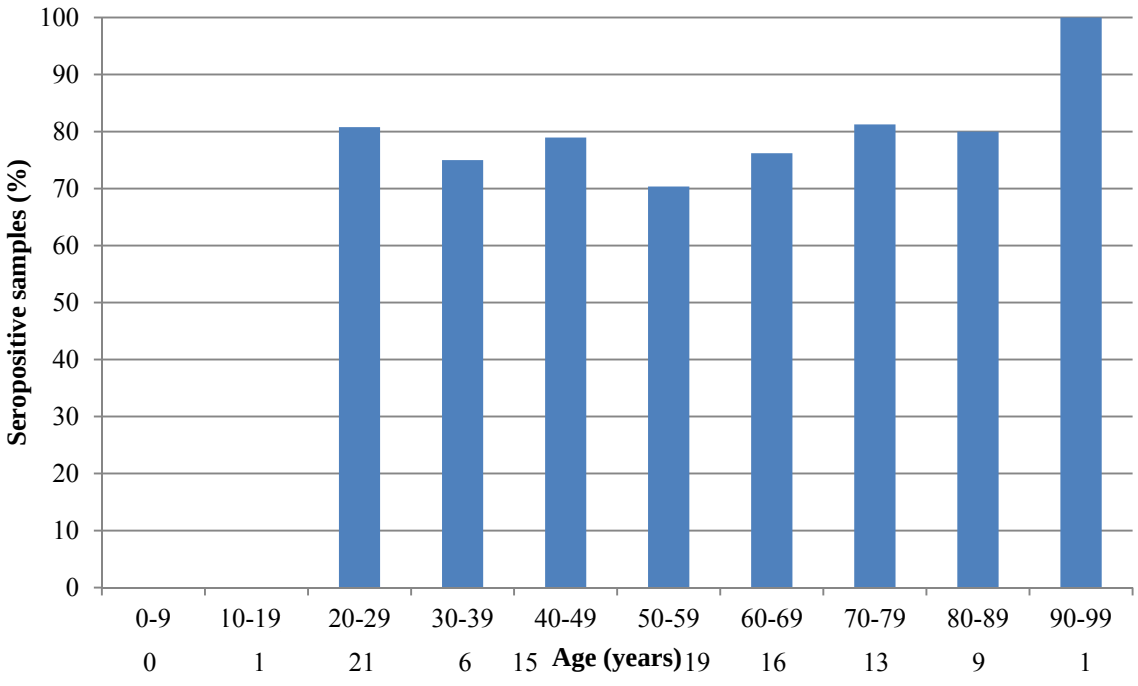


Figure 27 - FluC seropositivity in comparison to age

Due to the low number of participants in the 0-9, 10-19 and 90-99 groups, the results cannot be relied upon. The numbers below each of the groups on the x axis represents the number of participants in that group. Overall the figure shows the absence of a bell-shaped curve, and therefore indicates the data are not normally distributed and are therefore non parametric. Due to this, the data should be further analysed via Mann Whitney U tests.

Six different Mann Whitney U tests were done to fully compare the different groups. The comparisons were: male asymptomatic and female asymptomatic fluC IgG antibody levels; male symptomatic and female symptomatic fluC IgG antibody levels; male asymptomatic and female asymptomatic total IgG antibody levels; male symptomatic and female symptomatic total IgG antibody levels; symptomatic and asymptomatic fluC IgG antibody levels; symptomatic and asymptomatic total IgG antibody levels.

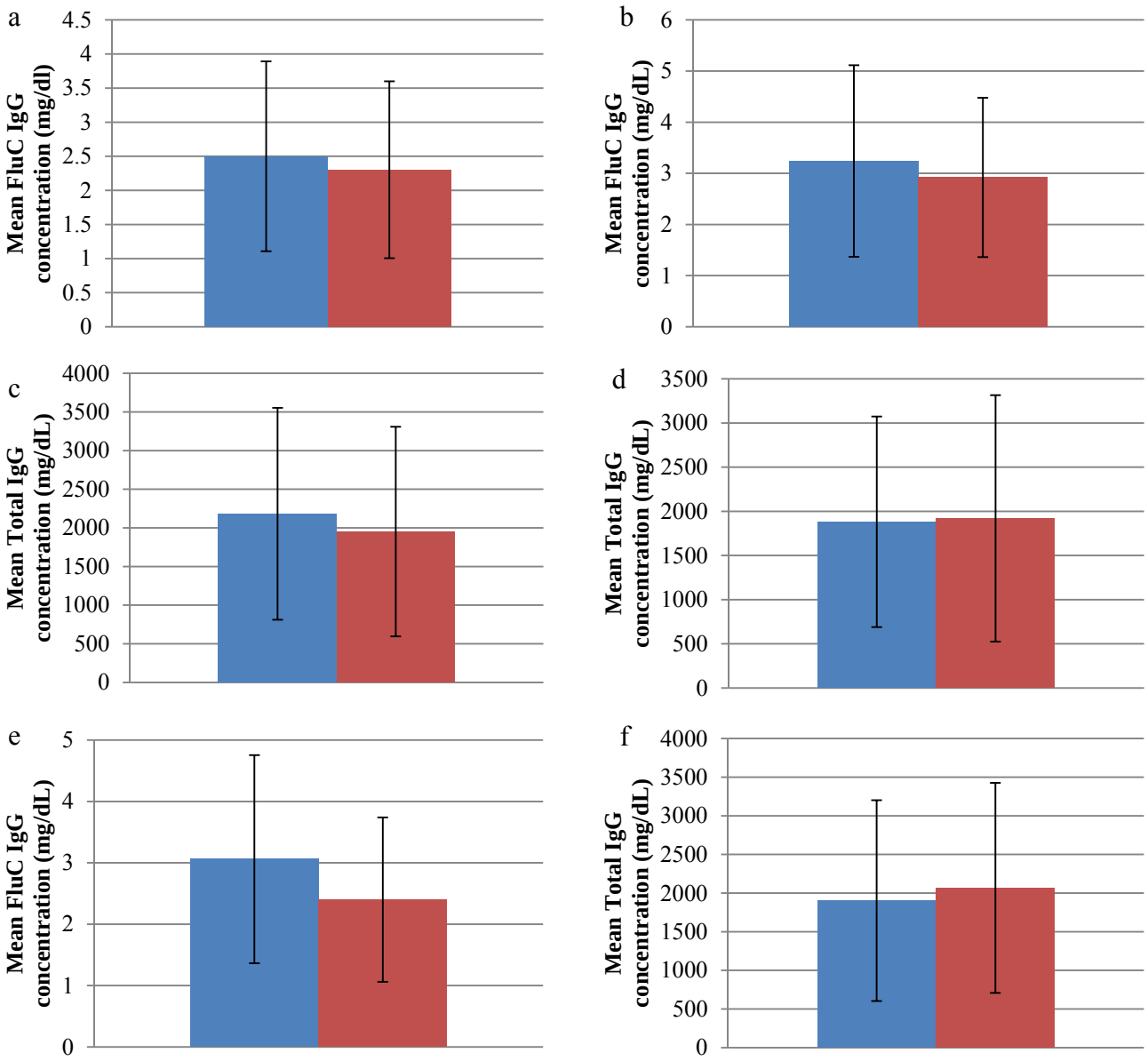


Figure 28 – Mean FluC and Total IgG Antibody levels for different population groups

A-f above compare the mean IgG levels for each group, with error bars added showing the variance of the data, via one standard deviation above and below the mean. a) Male control and female asymptomatic fluC IgG antibody levels. Blue indicates male asymptomatic and red indicates female asymptomatic. b) Male symptomatic and female symptomatic fluC IgG antibody levels. Blue indicates male symptomatic and red indicates female symptomatic. c) Male asymptomatic and female asymptomatic total IgG antibody levels. Blue indicates male asymptomatic and red indicates female asymptomatic. d) Male symptomatic and female

symptomatic total IgG antibody levels. Blue indicates male symptomatic and red indicates female symptomatic. e) Symptomatic and asymptomatic fluC IgG antibody levels. Blue indicates symptomatic and red indicates asymptomatic. This is the only set of results which is significantly different. f) Symptomatic and asymptomatic total IgG antibody levels. Blue indicates symptomatic and red indicates asymptomatic

Table 14 - Mann Whitney U Test

This table shows the z value results for each of the six groups analysed using the Mann Whitney U tests. The calculated z values were analysed using a standard table of z values in order to determine which groups were statistically significant, and if so at what parameters. There is only one significant test – the symptomatic fluC IgG levels and the asymptomatic fluC IgG levels ($p<0.05$).

Control	Experimental	Z value	Significant?
Male Asymptomatic Flu C IgG	Female Asymptomatic Flu C IgG	0.608164	Not significant
Male Symptomatic Flu C IgG	Female Symptomatic Flu C IgG	0.498532	Not significant
Male Asymptomatic Total IgG	Male Asymptomatic Total IgG	0.687	Not significant
Male Symptomatic Total IgG	Female Symptomatic Total IgG	0.048245	Not significant
Symptomatic Flu C IgG	Asymptomatic Flu C IgG	2.205303	Significant at $P<0.05$, not significant at $P<0.01$
Symptomatic Total IgG	Asymptomatic Total IgG	0.602308	Not significant

III.IV Plate results

Table 15 shows the overall *Staphylococcus* plate results for all of the nasal and nasopharyngeal swab samples. 'N/A' indicates either that the participant did not provide a swab sample or that the swab was not streaked onto a *Staphylococcus* selective plate. Samples that grew bacteria were indicated as *Staphylococcus* present, and those which developed a clear halo around the bacterial colonies were marked as coagulase positive, therefore indicating that the bacteria are *Staphylococcus aureus*.

Table 15 - *Staphylococcus* plate results (see appendix 5)

Overall 91 swabs were positive for *S. aureus*, 37 swabs were positive for other forms of *Staphylococcus* and 12 swabs were negative.

Tables 16-22 show the *Staphylococcus* plate results for different population groups.

Table 16 - *Staphylococcus* plate results summary

		Staphylococcus		
		Positive	Negative	Total
Coagulase	Positive	91	0	91
	Negative	37	12	49
	Total	128	12	140

Total percentage *Staphylococcus* coagulase positive = 65% (*S. aureus* positive)

Table 17 – *Staphylococcus* plate results – Symptomatic

Symptomatic				
		Staphylococcus		
		Positive	Negative	Total
Coagulase	Positive	40	0	40
	Negative	20	6	26
	Total	60	6	66

Total percentage *Staphylococcus* coagulase positive = 61%

Table 18 – *Staphylococcus* plate results – Asymptomatic

Asymptomatic				
		Staphylococcus		
		Positive	Negative	Total
Coagulase	Positive	51	0	51
	Negative	17	6	23
	Total	68	6	74

Total percentage *Staphylococcus* coagulase positive = 69%

Table 19 - *Staphylococcus* plate results – Nasopharyngeal

Nasopharyngeal				
		Staphylococcus		
		Positive	Negative	Total
Coagulase	Positive	74	0	74
	Negative	30	11	41
	Total	104	11	115

Total percentage *Staphylococcus* coagulase positive = 64%

Table 20 – *Staphylococcus* plate results – Nasal

Nasal				
		Staphylococcus		
		Positive	Negative	Total
Coagulase	Positive	17	0	17
	Negative	7	1	8
	Total	24	1	25

Total percentage *Staphylococcus* coagulase positive = 68%

Table 21 – *Staphylococcus* plate results – Female

Female				
		Staphylococcus		
		Positive	Negative	Total
Coagulase	Positive	40	0	40
	Negative	22	12	34
	Total	62	12	74

Total percentage *Staphylococcus* coagulase positive = 54%

Table 22 – *Staphylococcus* plate results – Male

Male				
		Staphylococcus		
		Positive	Negative	Total
Coagulase	Positive	51	0	51
	Negative	15	0	15
	Total	66	0	66

Total percentage *Staphylococcus* coagulase positive = 77%

III.V Bioinformatics and phylogenetic analysis

III.V.I Reproduction values

Reproduction values, or basic reproduction numbers (R), can be used to estimate how contagious a virus is. At a time of '0', the R_0 value for a specific virus will demonstrate how many other hosts the infected host can expect to infect from that point onwards. For example an R_0 value of '2' would indicate that an infected individual would be expected to infect two people. The R values of all three influenza viruses were analysed using the BEAST (Bayesian Evolutionary Analysis Sampling Trees) software package v1.8.0 (95) and Tracer software v1.6 (98). In order to calculate these data, all of the available fluC strains for each segment were downloaded in 'FASTA' format from the NCBI Influenza Virus Resource (92), before being aligned by 'ClustalW' in the MEGA software programme (97). Using these aligned

strains the best model was then determined using MEGA, before transferring the results to other bioinformatics software. Finally, using the software, the best clock was determined for each segment and all of the calculated data was then entered into BEAST and Tracer software packages in order to produce the R_0 values.

Table 23 – FluC R_0 values, with best fitting model and clock

This table shows the R_0 values for each of the seven segments, as calculated using the best model and clock for each specific segment. Each segment was run in the software to determine which model and clock was the most appropriate for that segment. In order to accurately calculate the R_0 value for a segment, the appropriate model and clock for that segment needed to be used. The exponential (exp) and logarithmic (log) clocks are both relaxed clock models, as opposed to strict. The substitution models of best fit are either the general time-reversible model (GTR) or the Tamura-Nei, 93 model (TN93).

FluC segment	Model	Clock	R_0 values
PB2	TN93	exp	1.029
PB1	GTR	exp	1.031
P3	TN93	exp	1.031
HE	GTR	log	1.486
MP	GTR	log	1.068
NP	TN93	exp	1.101
NS	GTR	exp	1.033

This process was then applied to available fluA and fluB segments for comparison, and as mentioned previously, fluC is more closely related to fluB than fluA, and so it was to be expected that their R_0 values would be similar. Table 24 shows the R_0 value results for all of the segments for fluA, fluB and fluC. The data for fluB and C are both from human host populations, whereas the fluA data are from an avian host population. The mean R_0 values are as follows: fluA – 2.15, fluB – 1.17 and fluC - 1.11. As fluA has an R_0 value around twice as large as the R_0 values for fluB and fluC, it shows that fluA is the most contagious

virus, and almost twice as contagious as fluB and fluC. FluB appears to be slightly more contagious than fluC, but overall the two viruses have very similar values.

Table 24 - R_0 values – fluA, fluB and fluC

This table shows the R_0 value for each of the segments of the fluA, fluB and fluC viruses. The mean and standard deviation values are also shown. Overall fluA has the highest average R_0 value – 2.15

			R_0 value		
Influenza virus			A	B	C
Host			Avian	Human	Human
Segment	PB2		2.21	1.16	1.03
	PB1		2.29	1.12	1.03
	PA/P3		2.29	1.14	1.03
	HA/HE		2.02	1.07	1.49
	NP		2.27	1.11	1.07
	NA		2.36	1.12	-
	M1		1.50	1.46	1.10
	NS	NS1	2.39	1.11	1.03
		NS2	2.00	1.26	
Mean			2.15	1.17	1.11
Standard deviation (SD)			0.28	0.12	0.17
Lower range			1.87	1.06	0.94
Upper range			2.42	1.29	1.28
No. of values within 1 SD			8	9	6
Total no. of values			9	9	7

III.V.II Phylogenetic trees

Using various bioinformatics software, phylogenetic trees have been created for all seven segments of the fluC virus. All of the available strains in the NCBI Influenza Virus Database (92)/GenBank (101) were downloaded and all full genome segments were used to construct the trees. The phylogenetic trees allow for comparison between the different strains of the virus, so that conclusions can be drawn as to which strains are more closely related and which

strains have evolved from which others. Figures 29- 35 show the phylogenetic trees for each of the seven segments. The ‘x axis’ under each tree represents time and the point on the left of the tree represents a fixed point in time from which all of the various strains have evolved. Each tree therefore visually demonstrates the potential evolution of the different strains over time.

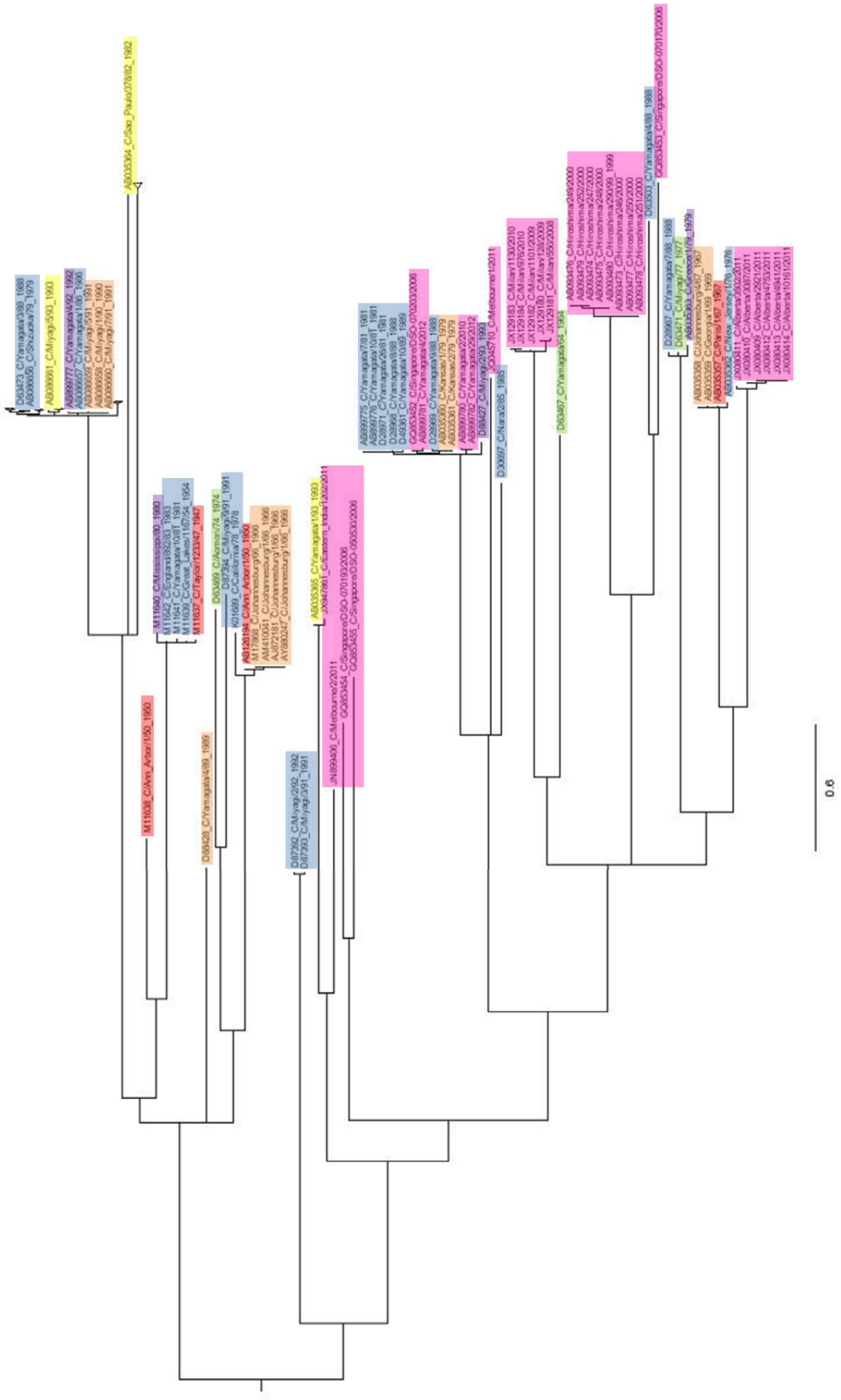

AB033064_C/Sao_Paulo/378/82_1982
D63473_C/Yamagata/3/88_1988
AB036656_C/Sn.zuoka/79_1979
AB000661_C/Miyagi/593_1993
AB000657_C/Yamagata/1/88_1988
AB000659_C/Miyagi/591_1991
AB000658_C/Miyagi/190_1990
AB000660_C/Miyagi/791_1991
M11640_C/Mississippi/80_1980
M11642_C/England/892/83_1983
M11641_C/Yamagata/108T_1981
M11639_C/Great_Lakes/1167/54_1954
M11637_C/Taylor/123/47_1947
D63469_C/Aomori/74_1974
D87394_C/Miyagi/191_1991
K01989_C/Beijing/1/79_1979
AB126184_C/Ann_Arbor/1/66_1966
M17808_C/Johannesburg/166_1996
AM410041_C/Johannesburg/166_1996
AJ872181_C/Johannesburg/166_1996
AY680247_C/Johannesburg/166_1996
M11688_C/Ann_Arbor/1/50_1950
D88428_C/Yamagata/4/89_1989
D87392_C/Miyagi/292_1992
D87393_C/Miyagi/391_1991
AB033065_C/Yamagata/1/93_1993
JX947661_C/Eastern_India/126/2011
JN899406_C/Melbourne/2/2011
GQ853454_C/SingaporeDSO-070193/2006
GQ853455_C/SingaporeDSO-050530/2006
A8399775_C/Yamagata/7/81_1981
A8399776_C/Yamagata/108T_1981
D29971_C/Yamagata/26/81_1981
D29968_C/Yamagata/8/88_1988
D49361_C/Yamagata/10/89_1989
GQ853452_C/SingaporeDSO-070203/2006
A8399781_C/Yamagata/4/2012
D29969_C/Yamagata/9/88_1988
AB033063_C/Yamagata/1/73_1973
AB033361_C/Kansai/2/79_1979
A8399780_C/Yamagata/2/2010
A8399782_C/Yamagata/29/2012
D88427_C/Miyagi/2/93_1993
D00697_C/Nara/2/85_1985
JO45710_C/Melbourne/1/2011
JX129183_C/Milan/113/2010
JX129184_C/Milan/99/2010
JX129185_C/Milan/110/2009
JX129187_C/Milan/116/2009
JX129181_C/Milan/55/2008
D63467_C/Yamagata/64_1964
A8393476_C/Hiroshima/24/2000
A8393479_C/Hiroshima/252/2000
A8393474_C/Hiroshima/247/2000
A8393475_C/Hiroshima/248/2000
A8393480_C/Hiroshima/250/99_1999
A8393473_C/Hiroshima/246/2000
A8393477_C/Hiroshima/251/2000
A8393478_C/Hiroshima/253/2000
D63503_C/Yamagata/4/88_1988
GQ853453_C/SingaporeDSO-070170/2006
D29967_C/Yamagata/7/88_1988
D63471_C/Miyagi/77_1977
AB033358_C/Johannesburg/467_1967
AB033359_C/Georgia/1/69_1969
AB033357_C/Paris/1/67_1967
AB033362_C/New_Jersey/7/76_1976
JX000411_C/Alberta/1/2011
JX000410_C/Alberta/3/2011
JX000412_C/Alberta/4753/2011
JX000409_C/Alberta/29212/2011
JX000413_C/Alberta/49412/2011
JX000414_C/Alberta/101612/2011
0.6

Figure 29 - Phylogenetic tree of fluC HE segment

This phylogenetic tree was created using the MEGA(97) and BEAST(95) software packages and was visualised using the FigTree software within BEAST (99). The 'x axis' represents time and the point on the left of the tree represents a fixed point in time from which all of the various strains have evolved. The tree prior used to create the tree was coalescent constant size.

The various branches have been colour coded in order to allow for comparison to Figure 6 by Matsuzaki *et al.* 2003) (31), reproduced overleaf. Each colour represents either one of the six lineages shown in Figure 6 - Taylor/47 (red), Kanagawa/1/76 (KA176) (green), Yamagata/26/81 (YA2681) (blue), Aichi/1/81 (AI181) (orange), Sao Paulo/378/82 (SP82) (yellow), and Mississippi/80 (MS80) (purple) or if the strain is a new strain, which has been found after Figure 6 was produced, then it has been coloured pink.

The Taylor/47 lineage contains both the Taylor/47 and Ann Arbor/50 strains, both of which can be seen in adjacent branches of Figure 29, in the second cluster from the top of the tree. Some of the strains from the Aichi/1/81 lineage can be seen along the bottom branches of Figure 29, such as Georgia/69 and Johannesburg/67, but some of the branches are also clustered towards the top branches of the tree. The Sao Paulo/82 lineage is also clearly visible as an isolated branch on the top right of the figure, however despite all three strains from this lineage being in the top half of the tree, they are not very well clustered. Again, the Yamagata/81 lineage is visible in Figure 29 below that of Taylor/47, but it is possible that these two lineages are not as clearly distinct in Figure 29 as they are in Figure 6. The Yamagata/81 lineage originally contained the most strains on the original Matsuzaki tree, and although this is still the case in this phylogenetic tree, the branches appear far more diverse. The remaining two lineages – Kanagawa/76 and Mississippi/80 – are not as clearly visible. Two of the strains within the Mississippi lineage Greece/79 and Nara/85 are in different

clusters of branches in Figure 29, suggesting that the phylogenetic tree in Figure 29 does not have as strong a suggestion of the six lineages present in Figure 6. All of the remaining strains coloured in pink are ‘new’ strains that were not available when Figure 6 was created. This increase in data may help to explain why the two trees in Figure 6 and Figure 29 appear different.

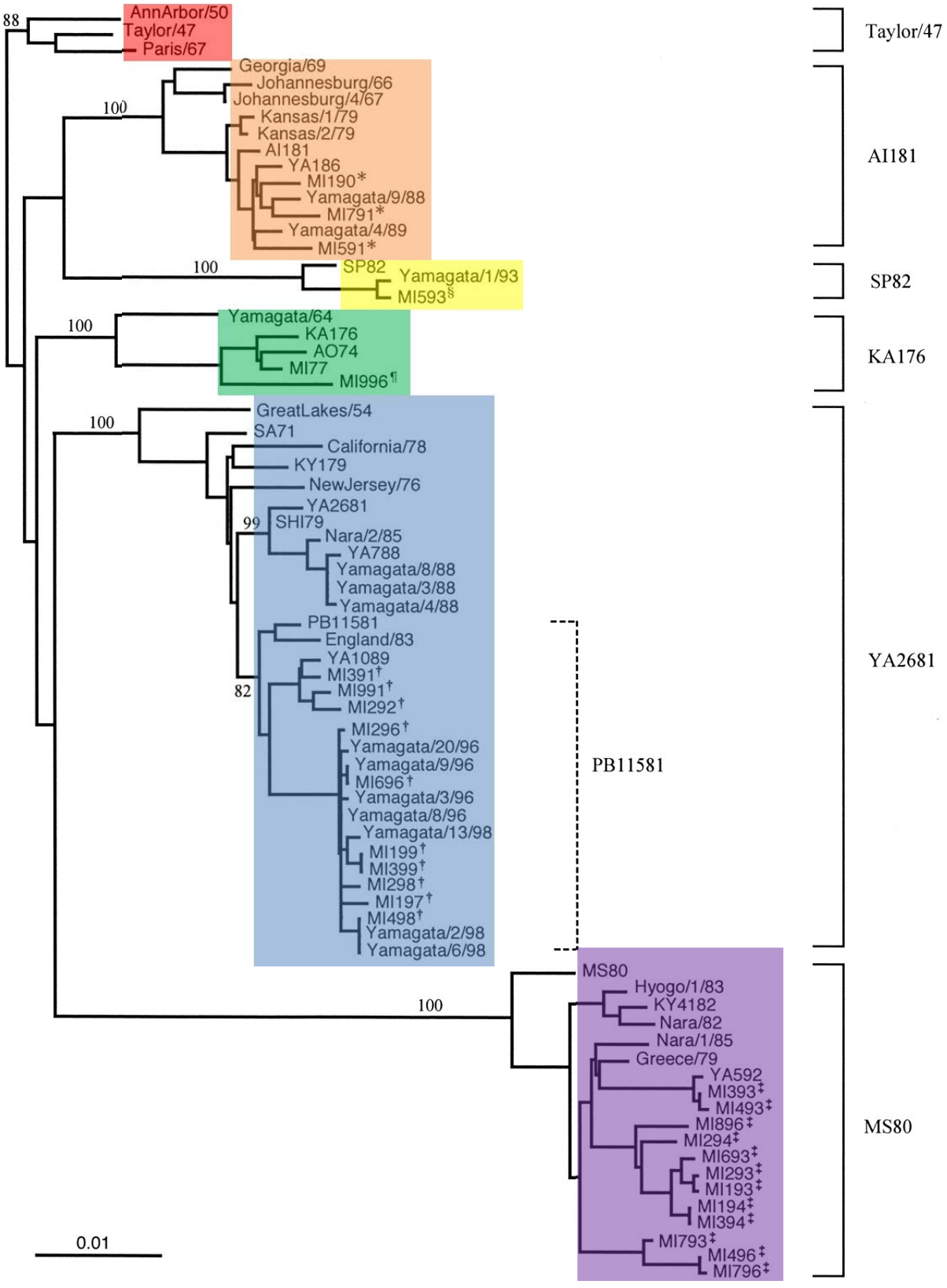


Figure 30 – (Repeated) Phylogenetic tree of fluC virus HE genes (Matsuzaki *et al.* 2003)

(31)

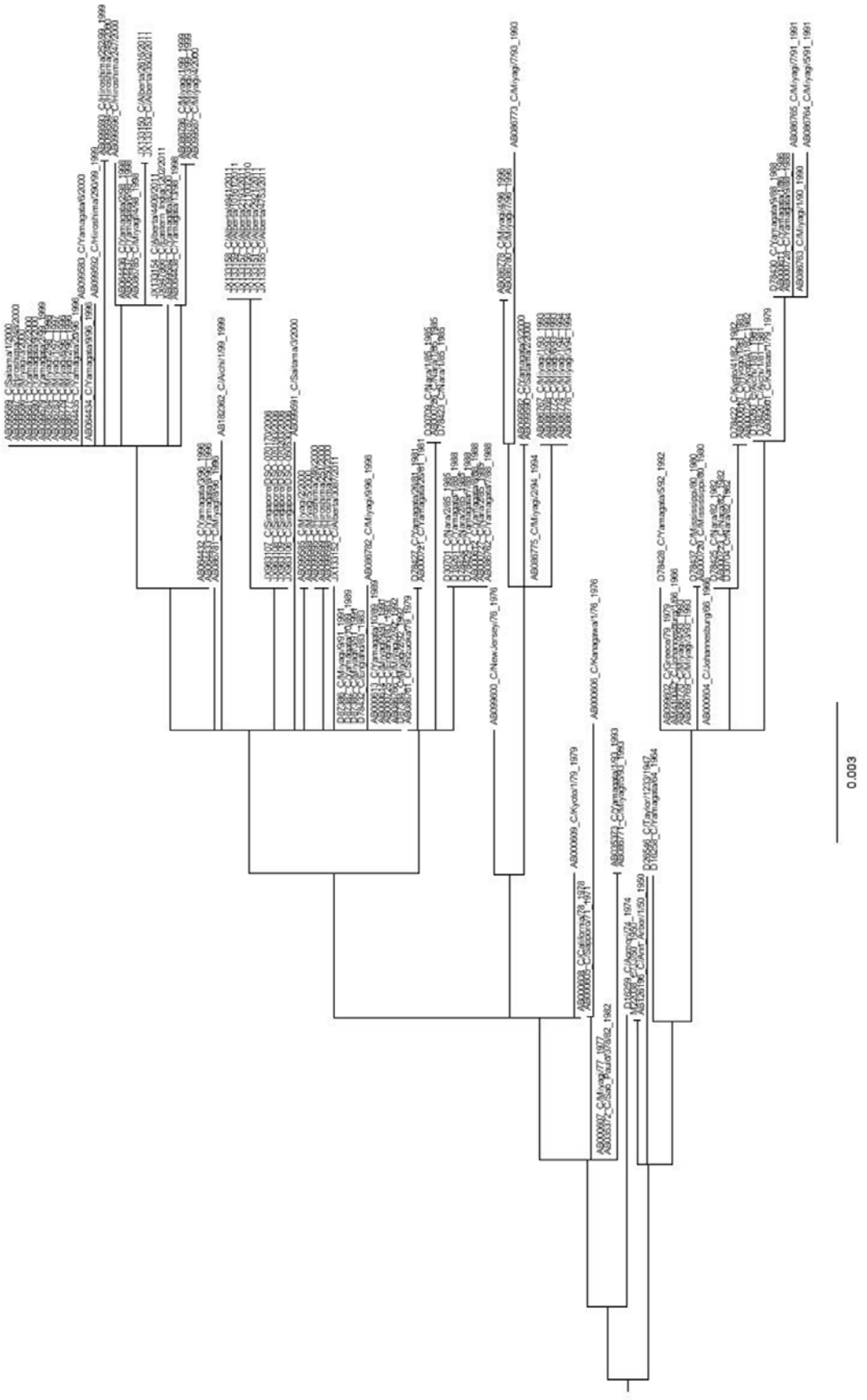


Figure 31 - Phylogenetic tree of fluC MP segment

This tree was produced using the same software and techniques as Figure 29.

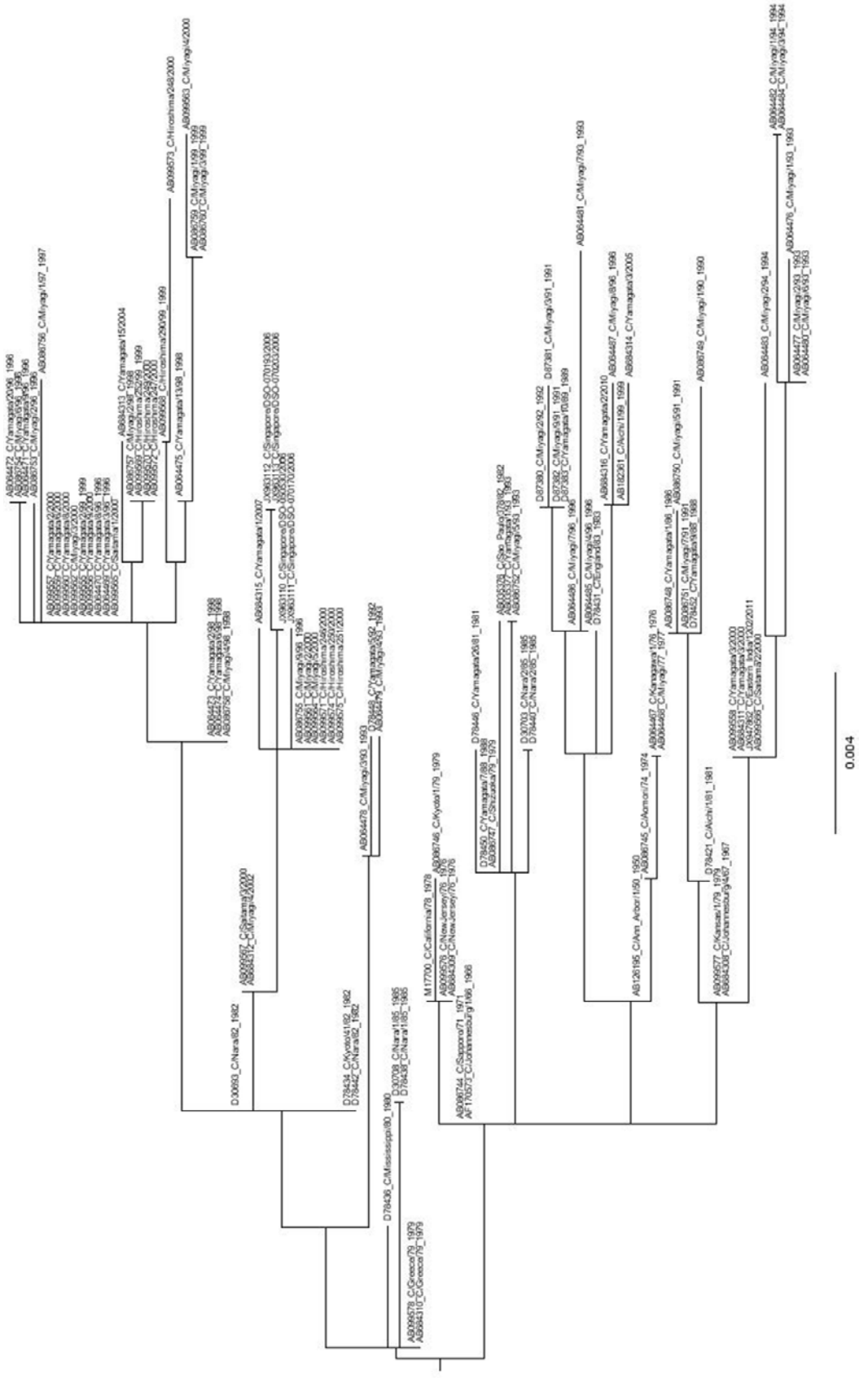


Figure 32 - Phylogenetic tree of fluC NP segment

This tree was produced using the same software and techniques as Figure 29.

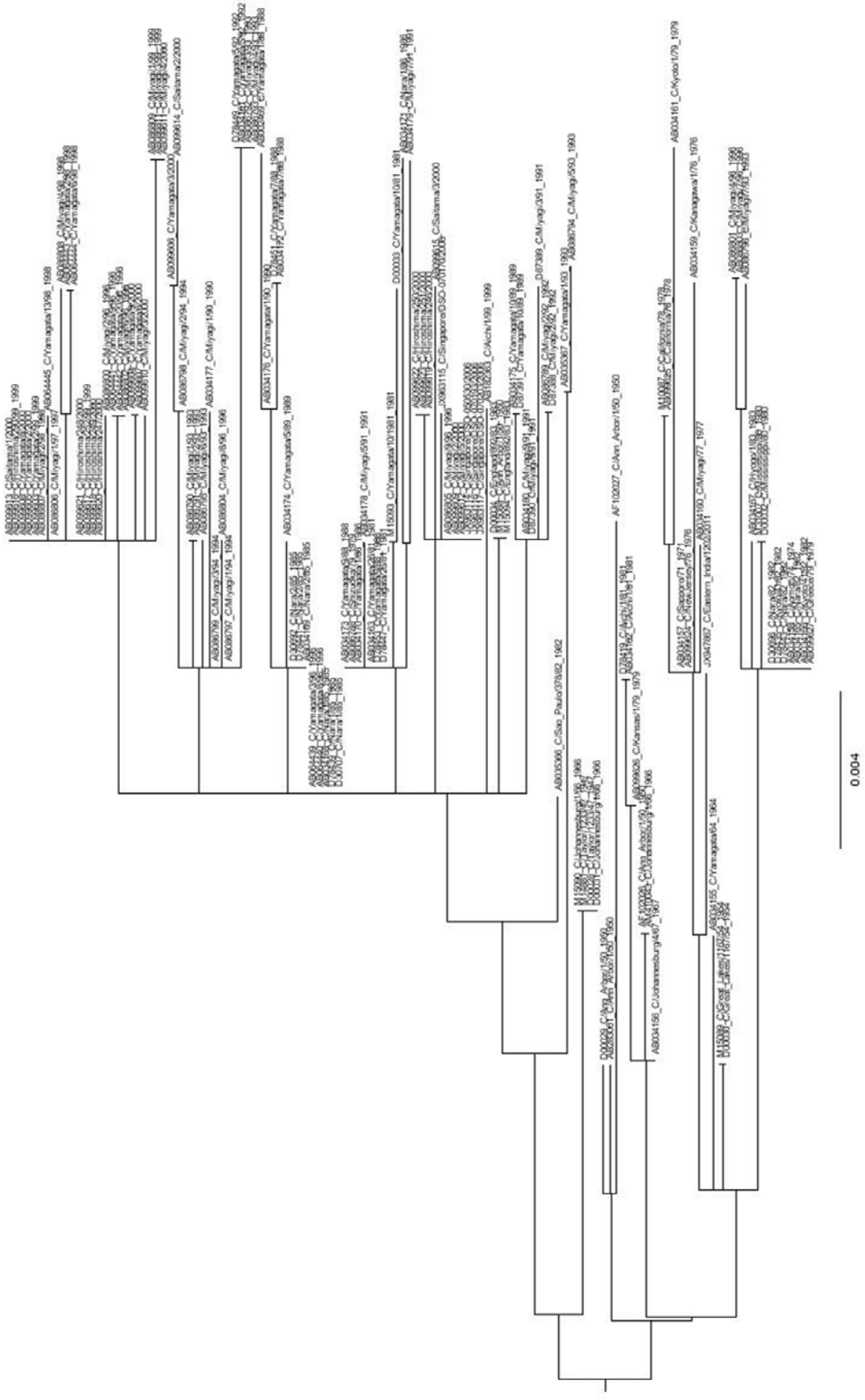


Figure 33 - Phylogenetic tree of fluC NS segment

This tree was produced using the same software and techniques as Figure 29.

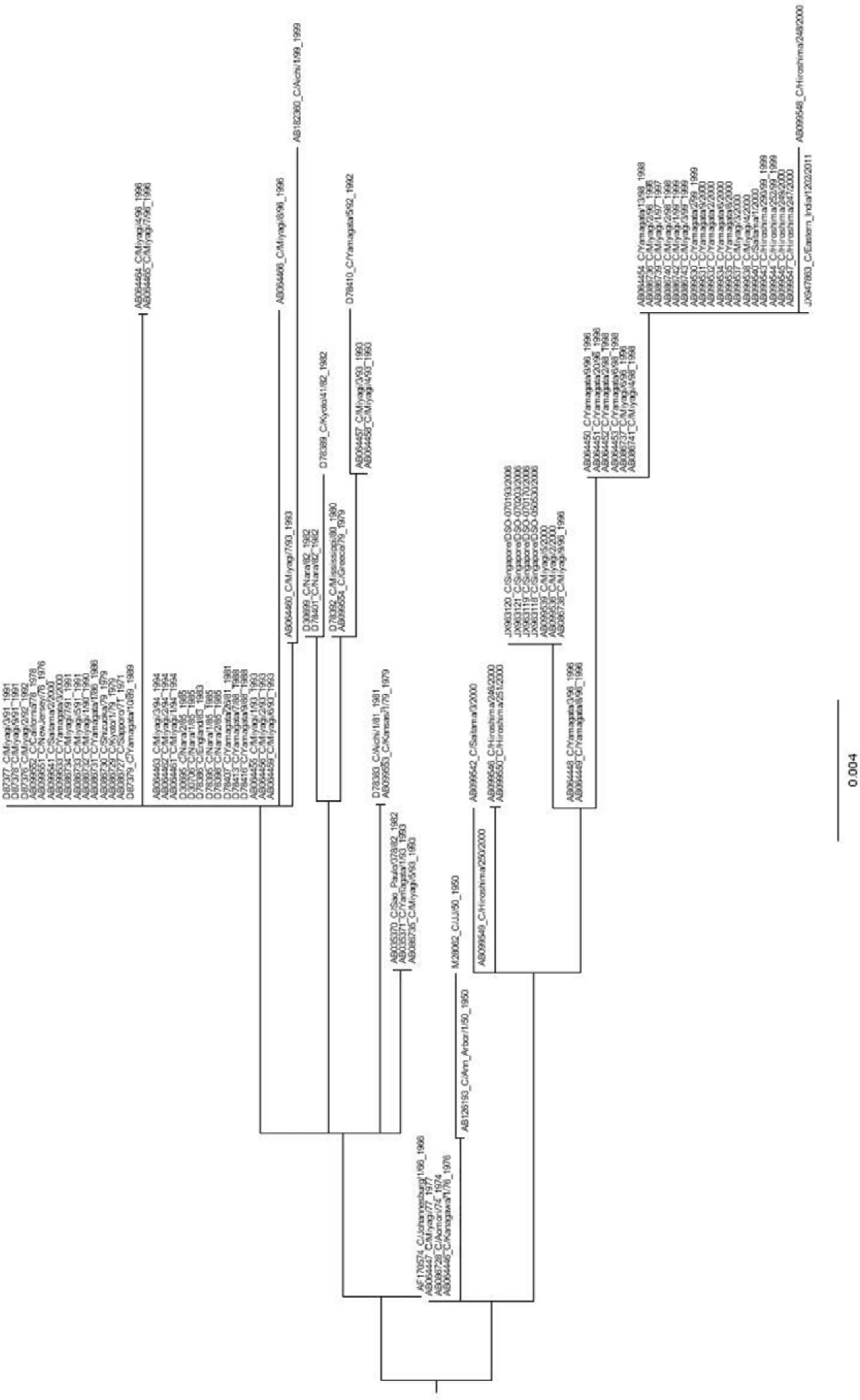


Figure 34 - Phylogenetic tree of fluC P3 segment

This tree was produced using the same software and techniques as Figure 29.

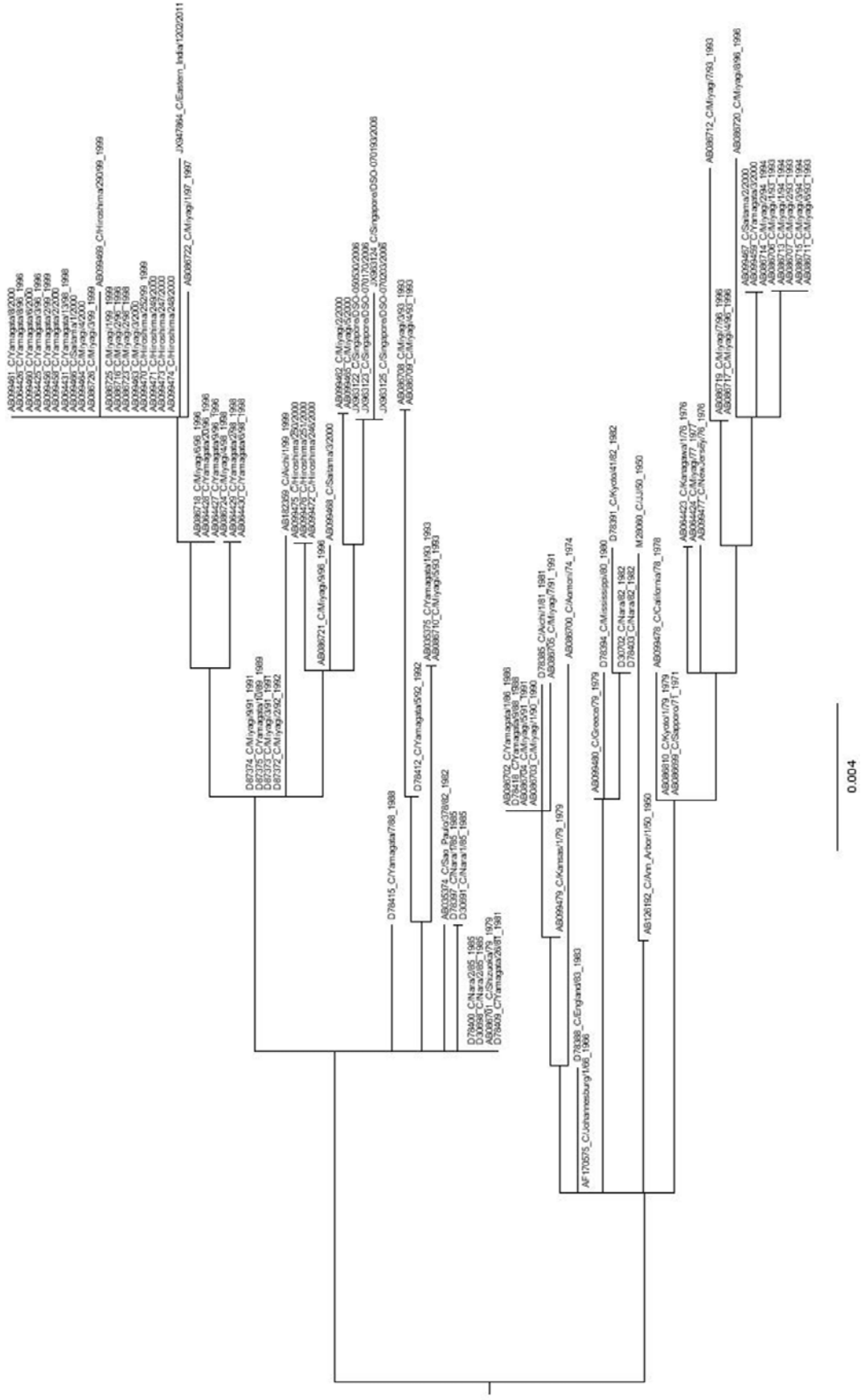


Figure 35 - Phylogenetic tree of fluC PB1 segment

This tree was produced using the same software and techniques as Figure 29.

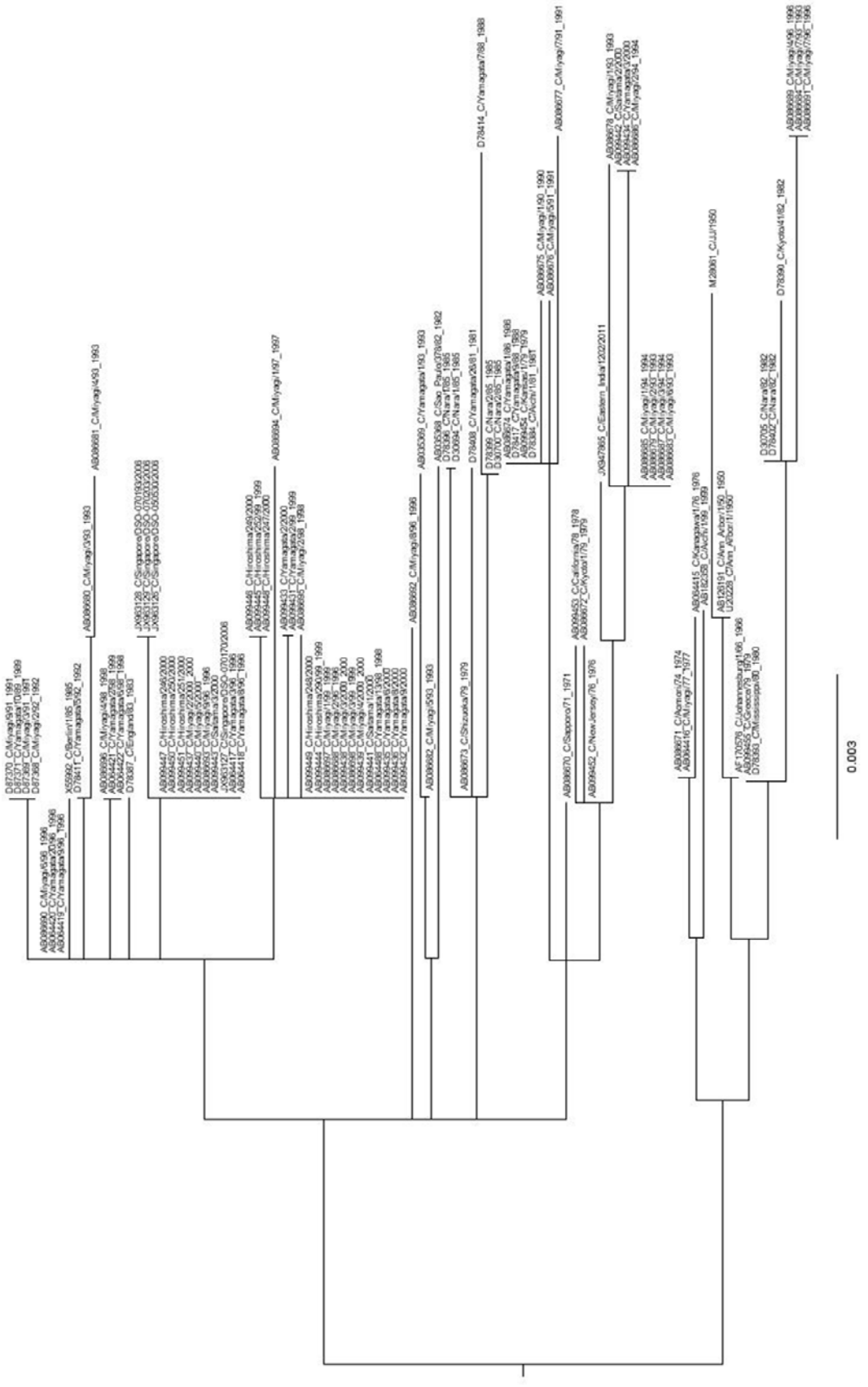


Figure 36 - Phylogenetic tree of fluC PB2 segment

This tree was produced using the same software and techniques as Figure 29.

IV Discussion

The main aims of this research were to determine whether the fluC virus is prevalent in Lancaster and whether the general population have antibodies against the virus. In order to meet these aims, serum samples and nasal/nasopharyngeal swab samples had to be collected and analysed using PCRs and ELISAs. All data then had to be analysed statistically and compared to current literature in order to be able to draw conclusions about the significance of the work and results.

IV.I PCR

Despite there already being set PCR protocols and primers created for the detection of fluC, various issues arose throughout various stages of the PCR process. Firstly when isolating the RNA, it became apparent when looking at the NanoDrop (ND) results that the carrier RNA supplied with the kit could potentially be masking any RNA isolated from the samples. In order to rectify this, linear acrylamide was purchased and was used in direct replacement of the carrier RNA.

When checking ND values for the samples there were also issues raised about the quality of the samples and whether any RNA had been extracted. Initially the 'blank' control sample used was water, which resulted in some samples obtaining a negative ND value. The blank reference value was then changed to a blank swab (a sterile swab which had been put through the same RNA extraction process), which gave better RNA values. Despite this, some ND values were still negative; however ND values ranged from -8.0 to 118.5 ng/µl. According to the manufacturer's instructions, the machine can accurately detect RNA in samples above 2 ng/µl, but as 37 samples had a negative ND value, it seems unlikely that these values can be relied upon. The positive control was also analysed via ND, and came out with a value of -32 ng/µl, the lowest value out of all of the samples. As it has been proven that this sample works

during PCR and therefore contains RNA, it also supports the view that the ND scores aren't reliable.

Another issue during the RNA extraction phase was caused by the nasal swabs (Copan). These swabs came with a sponge for transportation and preservation rather than a solution. As the nasopharyngeal swabs came with a viral transport medium, the solution could then be used immediately with the RNA isolation kit. As the nasal swabs did not, the samples had to be resuspended in solution in order to extract the sample from the swab tip.

In order to have the best chance of detecting any positive samples, the PCR had to be optimised. Performing the qPCR required the use of a master mix, and so the TaqMan® Gene Expression Master Mix (Applied Biosystems®, Life Technologies™) was purchased. The positive control provided by Nicolas Salez was claimed to have a CT value of 28, however when tested using the Gene Expression Master Mix this was not the case. A second qPCR was then trialled using this master mix and comparing it to a 'Universal Master Mix' that was already in the lab. The universal master mix gave a better CT value and so the samples were initially analysed using this master mix. After testing two full plates of samples, there were no positive results and some strange artefacts on the readings, and so an inhibition test was performed. The inhibition test showed the positive control samples to have worked, but some samples gave higher positive readings than the positive control. This would suggest that there is no inhibiting factor in the master mix or other element of the solution and would also appear as if some of the samples may have either been positive or something in the mixture actually facilitates the reaction. As they results were slightly inconclusive it was decided that a third master mix would be tested to see if that provided more reliable results.

A test qPCR was run comparing the Universal Master Mix with the TaqMan® Environmental Master Mix (EMM). The EMM gave a lower CT value and so it was decided to use the EMM to test the samples. The ambiguous samples that were previously tested were repeated; the

positive and negative controls worked well and the artefacts were no longer present on the results, showing that the plate had now worked. The remaining samples were also analysed and it was found that two samples came up positive, as shown in Figure 11.

IV.I.I FluC prevalence

Out of 148 nasal and nasopharyngeal swab samples tested, two samples (1.35%) were detected as being positive via PCR. It was expected that any positive samples would appear in participants who had typical 'flu' symptoms, or those with symptoms that have previously been described in a fluC infection. Surprisingly, both positive samples were found in asymptomatic patients who claimed that they were asymptomatic. This could be due to a number of reasons; perhaps the participants simply had strong immune systems and were therefore asymptomatic, or the strain that they were infected with was mild and so they had not chosen to complain about their symptoms. The participants could also currently be in the incubation stage of the virus and although were carrying the virus were yet to exhibit symptoms. In contrast to this, they could have been recovering from a recent infection and failed to mention when asked that they had had a recent infection and were therefore included in the asymptomatic group.

As the positive samples were collected from potentially asymptomatic carriers, it is therefore difficult to compare the statistics in this study to those found in other studies. Other papers have found positive results in symptomatic patients and have been able to compare symptoms etc. and other clinical factors, but that is not possible in this case.

When looking back at the data, both positive swabs were collected a day apart at the end of January 2015. This matches with most other papers, which suggest that the peak season for fluC is January-June (23, 25-27, 29-35). The first samples were collected in November 2014

and the very last sample was collected in May 2015, however the majority of samples were collected in January (93 samples) and February (39 samples) 2015.

IV.I.II Deep sequencing

One of the large decisions to make was which samples were to be sent for deep sequencing. Ideally, the samples sent needed to optimise the chance of finding a positive fluC sample in case any were missed via PCR. The samples sent also needed to be varied enough that comparisons could potentially be made between groups of samples. Originally four mixtures were sent – one containing the paediatric samples, one containing symptomatic adult samples with high fluC IgG levels, another contained all of the samples from adults with pyrexia and finally there was an asymptomatic group containing asymptomatic adult samples with low IgG levels.

Unfortunately it was found that the samples that had been sent were single-stranded cDNA, rather than the double-stranded cDNA required for the deep sequencing. During this time two samples had also appeared positive on PCR and so needed to also be sent for deep sequencing.

A new cDNA synthesis kit was ordered which produced double-stranded cDNA; however this could only convert a maximum of 10 RNA samples to cDNA (and also required the purchase of the oligo (dT) primer which delayed matters further). It was decided to send nine samples for deep sequencing; however for some of these samples the RNA was pooled prior to conversion to cDNA in order to increase the diversity in the samples. Nine new groups were chosen (see section II.VII), with the aim being to create as vast a difference between the samples as possible, whilst still aiming to optimise the chances of retrieving fluC sequences. As fluC is a respiratory virus, the addition of the ‘asthma’ and COPD groups will allow for a

greater analysis of the impact of the virus, once the deep sequencing results have been returned.

IV.II Total IgG ELISA

Prior to testing the serum samples, the ELISA had to be optimised and several different methods were trialled prior to finding a successful protocol.

From all of the total IgG serum results, the mean was 1994.66 mg/dl (range 293.39- 5504.65 mg/dl). This is significantly higher than the reference values used of 994 mg/dl (639–1,349) (53, 54). This could be due to the estimation of the IgG levels by plotting them on the curves. The method is not exact as the curve is produced by using a slightly different method to the samples. There is also some discrepancy when choosing which points to use to draw the 'straight' aspect of the curve; the points are chosen by eye, based upon the viewer's opinion as to where the straight line starts. This could potentially mean two different people could choose different points to base the straight line on, therefore giving different results. This may help to partially explain why the total IgG levels found in this study are much more variable than in the literature. As the same process was used to determine fluC IgG levels, then the total IgG values still serve a purpose as they allow for a basic comparison between samples, and the total IgG values can also be used to check samples for any immunosuppression etc.. According to the literature, a few of the samples in this study would be classed as having an IgG deficiency (55, 56), but as stated by Puissant-Lubrano *et al*. many of their patients with decreased levels of IgG were actually healthy (56), and so it is difficult to draw conclusions on the participants in this study.

IV.III FluC IgG ELISA

In order to make sure the protocol sent by Salez would work, it was trialled and compared to a standard protocol already in use at Lancaster University for AH-specific IgG. The

university protocol worked well, however Salez's protocol did not work when applied to AH. As limited reagents were sent, the protocol could not be practiced too much as it risked using up the sera etc. The Lancaster protocol was practised, but adapted to be as close to Salez's protocol as possible, and was then applied to every venous serum sample to give an AH IgG OD value. A plate was then trialled following Salez's protocol exactly and testing only a couple of samples. This was successful, suggesting that the ELISA is very specific for fluC, and so all of the samples were then able to be tested.

IV.III.I FluC antibodies

As a similar study had not been carried out in this area before, there was uncertainty as to whether the general population would have antibodies to the virus, particularly as there has rarely been any documentation of fluC in the UK. The levels of seropositivity were higher than expected, but show that the fluC virus is in the area and that the study was worthwhile.

The mean fluC IgG concentration was 2.40 mg/dl (range 0.24 mg/dl – 5.88 mg/dl), which is around 10 times lower than that listed for fluA (24 mg/dl (range 15 – 39 mg/dl) (55)). This could be due to a number of reasons, the main one being that they are different viruses and it would be expected that fluA is far more prevalent, however different populations and detection techniques will also impact on the final values.

Statistics for the UK seroprevalence rates are around 60-70% (28, 70), and although the values from this study are slightly higher than may be expected, they are still in keeping with worldwide statistics (23, 27, 28, 32, 35). There are also discrepancies as to how to decide which samples to class as positive and negative. Using the original method of 1.96SD gave a seropositivity rate of 77%. This value, although slightly lower, supports the seropositivity rate calculated above and uses a similar method to Salez.

In order to be able to analyse the results between the different population groups, statistical tests were applied to the data. A chi square test was performed, comparing the male and female symptomatic and asymptomatic groups, and the seropositive results were compared to age, which showed that the data were both randomly allocated and not normally distributed. This then meant that the data had to be analysed using Mann Whitney U tests. Six different Mann Whitney U tests were then performed comparing six different groups of data; male and female asymptomatic fluC IgG, male and female symptomatic fluC IgG, male and female asymptomatic total IgG, symptomatic male and female total IgG, symptomatic and asymptomatic fluC IgG and symptomatic and asymptomatic total IgG. The only test which was significantly different at $p<0.05$ was the comparison between the symptomatic and asymptomatic and fluC IgG levels, which shows that the symptomatic participants have statistically significant higher levels of fluC IgG antibodies. This also shows that there is no difference between the male and female groups in any test and that the symptomatic and asymptomatic groups do not have significantly different total IgG levels.

IV.III.II Determining positive and negative serum samples

One area of discrepancy was the method used to determine which serum samples to class as positive and negative. Although every paper that performs ELISAs for this purpose must use a specific method or calculation, there is often no mention of this in the published articles.

It is statistically accepted that 95% of the data falls within 1.96 standard deviations (SD) of the mean. Based upon this principal, originally the data which fell within 1.96SD of the positive control mean were classed as positive, and those which fell within 1.96SD of the negative control mean were classed as negative. This gave a seropositivity rate of 77% which, although in keeping with the current literature, left 26 samples in a ‘grey area’ between the two groups of positive and negative samples.

A paper was then found testing serum samples via ELISA IgG antibodies to the Ebola virus (104). Although looking at a different virus, the principals were still the same and so their method was used to determine which samples to class as positive. Any sample which had an optical density value greater than 3SD above the mean negative control was classed as positive (104), which when applied to this data made only five samples out of 129 negative, giving a seropositivity rate of 96%. This figure is higher than all of the values in the literature and so may be unreliable.

The alternative method used was to compare our samples to those analysed by Salez *et al.* As the antigen and positive and negative control sera used in this study was provided by Salez, it was justifiable to compare the data from this study to that study. Salez worked out the mean of the negative sera and then calculated 2SD above and below the mean. Any samples which fell within this range were classed as negative and all other samples were classed as positive. As they collected and analysed samples in a variety of places, all of their samples were compared to the 'control' plate in Marseille, which was also done to the samples from this study. Using this calculation gave a seropositivity rate of 82% for this data.

IV.III.III Analysis of the products provided by Salez

In order to make sure that the negative control was a true negative, and not just an IgG depleted sample, such as from an immunocompromised patient, the total IgG level needed to be quantified. The results are shown in Table 7 - the positive serum had a total IgG concentration of 1822.63 mg/dl and the negative control had a concentration of 2047.27 mg/dl. As the mean total IgG concentration for all of the serum samples was 1994.66 mg/dl, this shows that both the positive and negative control sera have normal IgG levels. This then proves that the fluC ELISA is looking for specific IgG antibodies, and is not simply testing overall IgG concentrations.

Part way through the testing of the serum samples the initial antigen solution ran out and so a second solution had to be used. In order to confirm that these two antigen solutions were the same, they had to be compared. To do this, the new antigen was diluted to concentrations of both 1:400 and 1:800. The OD values for the Sigma serum and the positive and negative controls were read and compared to those for the old antigen. The results were then plotted on a graph in order to provide a visual representation (Figure 25). As this graph shows, both antigens give very similar OD results, and as the T tests showed, one can be confident that there is not a statistical difference between the two antigen solutions. The values at 1:800 were almost exactly at the expected values, around half of those at 1:400, which again shows the reliability of the results.

It was also noted that the negative control serum did not give an OD value of '0', which some might expect would be the definition of a truly negative sample. A final test was carried out on the control sera to check whether the sera would stick to the wells in the plate, regardless of the presence of the antigen, which may explain the reasoning behind the raised negative value. The negative control serum gave a higher OD reading (0.245), when there was no antigen present, than both the positive control serum (0.177) and the background wells (0.108). This does then appear to partially explain the 'raised' negative control readings.

IV.IV Screening of swabs for *Staphylococci*

All of the swabs were streaked onto *S. aureus* selective plates in order to prove that the swabs had made contact with the lining of the nostril or nasopharynx. One hundred and forty out of the 143 samples collected were tested, as the first three swabs were collected prior to obtaining the material needed to pour the plates. Overall 91% of the swabs grew some form of *Staphylococcus* and 65% of the swabs were coagulase positive, confirming the presence of *S.aureus*. There were coagulase positive samples in all age groups – the youngest was six

years old and the oldest was 90 years old. This confirms that the swabs have been successful in obtaining samples. This result is higher than documented in other studies, most studies have much lower values of colonisation of ≤35% (82, 83, 105-109), however one paper does state a nasal colonisation rate of 52.3% (110). These differences in values could be due to multiple reasons; for example all of the aforementioned papers are analysing nasal swabs, whereas the majority of patients in this study have provided a nasopharyngeal swab, although in this study there is little difference between the two sample types. Different detection methods, populations and cultural differences may also all result in varying colonisation rates.

When comparing the different population groups and types of swabs there was little difference. Symptomatic participant swabs showed that 61% of participants were *S.aureus* positive, whereas the control group had 69% positivity. The nasopharyngeal swabs grew *S.aureus* in 64% of cases, whereas the nasal swabs were 68%. The main difference observed was between the female and male swabs; the female participants grew *S.aureus* on 54% of the swabs, whereas the men grew *S.aureus* on 77% of the swabs. This is, however, in keeping with the literature (84, 85).

Interestingly, despite there being 12 completely negative swabs out of the 140 tested, all 12 came from female participants, although there was no explanation for this in any of the literature. These negative swabs did come from females of all ages – the youngest was 11months old and the oldest was 75 years old – and there were the same number of negative swabs in both the symptomatic and control groups. There were more negative swabs in the nasopharyngeal group (11 negative) compared to the nasal group (one negative) but this was to be expected, as there is less mucus etc. in the nasopharynx and so the environment is less favourable for bacterial growth.

IV.V Limitations

Despite the overall success of the study, there were some limitations which mean that the project could be improved upon. Firstly, the ethical approval took a few extra months than expected to be granted, and it was therefore mid-November before any sample collection was allowed to take place. The equipment required for sample collection then did not arrive until the following January, resulting in further delays.

Once sample collection began, there was then the issue of which patients to recruit. Any patients with symptoms of ILI, as mentioned previously, and who gave informed consent were included. It can be argued, however, that a person's experience and tolerance of flu-like symptoms can be quite subjective. Most patients would not choose to bother their GP with symptoms of a cold, and so it could be perceived that the patients recruited with the ILI symptoms could all be of a similar personality trait, and therefore do not reflect the entire population. It can also be viewed that the patients able to see their GP during standard working hours are mostly retired or unemployed, with only those who are too sick to work able to attend a daytime appointment. There is the alternative viewpoint that those recruited with ILI would be representative of the 'fluC population', as it has been mentioned before that the children and the elderly are most at risk (34, 47, 48).

Some inpatients from the RLI were recruited, but due to time pressures only nine people out of the 147 participants recruited were current inpatients on the respiratory ward. Ideally, it would have been a more robust study if more unwell patients were recruited, as this would have allowed for stronger conclusions to be drawn. As previous studies have mentioned pneumonias and serious complications, it would have been good to recruit more patients with this type of diagnosis. In addition to this, if more children had been recruited there would have been a greater chance of finding the live virus. Despite having permission and approval

to collect samples from the paediatric ward at the RLI, unfortunately no patients were suitable to participate in the study during the recruitment period.

One aspect of the fluC virus that would have been interesting to investigate is whether the virus is seasonal, or whether it is present in the population all year round. Due to time constraints again, samples were mostly collected in January and February, preventing any conclusions from being drawn about when the virus is most prevalent. There is a possibility that the main 'fluC season' could have been missed without our knowledge.

When choosing which primers to use for the PCR, there was a limit in the number of papers who have used this process. It was agreed to use the primers from the most recent paper published at the time (28), and also as this was from a nearby country, there was a greater chance of these primers matching any potential fluC viruses in England. One drawback to this is that other strains of the fluC virus may have been missed. All available M segments for the fluC strains were downloaded from GenBank (101) and were aligned using the MEGA software (97). The primers by Salez were searched for, and despite a good overall finding, not every strain had the specific primer sequences. This then suggests the possibility that some samples that were tested during this study could potentially have been positive for fluC, but if it was a different viral strain then it may not have appeared positive on PCR.

Other issues relating to the PCR arose. Thirty seven samples had negative ND values, which would suggest that there was no RNA in the samples at all, so the reliability of either the quality of the samples or the use of the ND machine needs to be questioned. The nasal swabs had to be resuspended in PBS as the swabs did not contain any transport medium. As none of these samples came out positive on PCR it could appear as if no RNA was extracted from the samples, even though the samples actually all had positive ND values.

Another issue with the PCR is that it may not have been optimised. The positive control is supposed to have a CT value of 28, but even when comparing three different master mixes the CT value still came out above 30. This could suggest that if a sample had a small amount of the RNA in it, it could take more than the 45 cycles of the PCR to be amplified and detected, possibly causing positive samples to appear negative.

The samples prepared for deep sequencing were made using a process which creates a very small white pellet. In all but one of the samples this pellet was not visible, and so there is a chance that either there was no RNA in the samples and therefore no pellet, or that the pellet could have accidentally been discarded with the supernatant. The pellet then had to be resuspended in 3µl of solution, which is also a very small amount to be frozen and transported for deep sequencing, so again there is a chance that it will appear as if there is nothing in the sample. Also due to time constraints, the deep sequencing results were unable to be analysed, however further work will be done to interpret the results.

The final limitations to the study were in relation to the analysis of the ELISA results. Positive and negative control sera were sent by Nicolas Salez for use in the ELISAs. Despite there being a big difference between the readings for the two controls, the negative control could be perceived as not being a true negative, as it had a much higher reading than the 'blank' wells. When no antigen was stuck down in the wells containing serum samples, the negative reading still gave a higher reading than the background wells without serum samples, suggesting that there might be some non-specific binding.

In order to estimate the IgG concentrations in each sample, the standard concentration curves were used. Based upon the straight part of the curve, an estimation of the equation of this part of the curve was used in order to calculate the approximate concentration of IgG in each sample. This can be argued as a matter of personal preference as to where to choose the points for the 'straight' aspect, and so in order to avoid this problem the same points were

used for each plate. This, however, causes its own problems, as each curve differs, and hence the values could be deemed as inaccurate. This also becomes clearer when looking at the total IgG concentrations for the samples; most samples fall either above or below the normal range due to the rough calculations and estimations used.

Further issues arose when deciding which samples to class as positive and negative. Using Salez's method gave a seropositivity rate of 82%, whereas using the 1.96 SD method gave a seropositivity rate of 77%, which although both give similar values, they could be interpreted differently. When reading the literature, most articles will not mention how they calculated which samples were positive and negative, despite this process taking place. It seems to be a process without a clear and agreed upon method, and so depending upon which method used the results will vary. This could potentially mean that the results from this study either over or under estimate the number of samples which were positive.

IV.VI Future

This study has provided a good foundation for future work; however there are many areas which can be built upon. Firstly, in order to get a good impression of the seasonality of fluC, samples would need to be collected all year round. Ideally if this could be continued over a few years, then it could be determined whether the virus is around in certain seasons, all year round, or simply has random epidemics that cannot be predicted.

Despite a reasonable number of people being recruited for this study, there was an uneven range of age groups, as more patients were recruited in the 20-29 and 50-59 age groups. If this project was to be taken forward, more patients would need to be recruited from other age groups in order to get a more accurate picture of antibody levels etc. in the general population. In particular, more focus needs to be placed on obtaining paediatric samples, as it

is children who are both most likely to become infected with the virus and to have more severe symptoms.

Further emphasis should also be placed on recruiting those with respiratory conditions. As fluC is a respiratory virus, it would seem logical that those with pre-existing lung conditions would be more susceptible, or would be more severely affected, by a fluC infection.

Finally, the use of multiple primers when performing PCR would decrease any concerns of missing 'positive' samples, due to the primer sequence not being present in that specific strain of the virus.

In the long term, routine monitoring of fluC would be ideal, as the gaps in documentation of the virus mean that one can only guess at the behaviour of the virus, rather than actively following it and predicting the next development, as is the case with fluA. If it was found that fluC is as troublesome as fluA and fluB, then an end goal would obviously be the inclusion of fluC in the annual influenza vaccine.

V Conclusion

FluC is a respiratory virus which can cause symptoms varying from mild colds to pneumonias with potentially serious complications, particularly in children. Despite this, few published articles show studies conducted on this virus and it is not routinely screened for. Older papers suggested that fluC would, at best, would cause mild 'cold' symptoms but this now appears not to be the case.

This study aimed to investigate whether fluC is prevalent in Lancaster and whether the general population have antibodies against the virus. In order to test these aims, nasal/nasopharyngeal swabs and serum samples were collected from 148 participants. PCR was then performed on the swab samples and the serum samples were analysed using ELISA.

It was found that every participant had some level of antibody to the virus, despite not all serum samples being classed as positive, therefore showing that the fluC virus is prevalent in Lancaster and that everyone has been exposed to the virus. Two samples also appeared positive on PCR for the fluC virus but, interestingly, both of these samples were in asymptomatic participants.

Due to limited time, not enough participants were recruited to fully analyse and compare the average antibody levels in the population etc. It would be of interest to see whether antibody levels remain constant into old age, as they do in Japan, or whether levels decrease, as appears to be the case in the rest of the world. In order to further strengthen this study, future work would also need to recruit people throughout the year to allow for further understanding of the seasonality of the virus.

Overall, this study has met its aims and has been successful in detecting both live virus itself and antibodies to fluC, showing that the virus is prevalent in Lancaster and providing further evidence for the need to study fluC in the future.

VI References

1. Taylor RM. Studies on survival of influenza virus between epidemics and antigenic variants of the virus. American Journal of Public Health and the Nations Health. 1949;39(2):171-8.

2. Encyclopaedia Britannica. Orthomyxovirus. 2014.

3. Maclachlan NJ, Dubovi EJ, Fenner F. Fenner's Veterinary Virology. 4th ed. / edited by N. Maclachlan and Edward J. Dubovi. Ed. London: Academic; 2011.

4. Collin EA, Sheng Z, Lang Y, Ma W, Hause BM, Li F. Cocirculation of two distinct genetic and antigenic lineages of proposed influenza D virus in cattle. Journal of Virology. 2015;89(2):1036-42.

5. Tidona CA, Darai G, Osmond CB. The Springer Index of Viruses. Berlin ; London: Springer; 2002.

6. Muraki Y, Hongo S, Sugawara K, Kitame F, Nakamura K. Evolution of the haemagglutinin-esterase gene of influenza C virus. The Journal of General Virology. 1996;77 (Pt 4):673-9.

7. Muraki Y, Hongo S. The molecular virology and reverse genetics of influenza C virus. Japanese Journal of Infectious Diseases. 2010;63(3):157-65.

8. PBD ID: 1FLC

X-Ray Structure of the Haemagglutinin-Esterase-Fusion Glycoprotein of Influenza C Virus. Rosenthal PB, Zhang XD, Formanowski F, Fitz W, Wong CH, Meier-Ewert H, Skehel JJ, Wiley DC. Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. Nature. 1998;396(6706):92-6.

9. Rosenthal PB, Zhang XD, Formanowski F, Fitz W, Wong CH, Meier-Ewert H, Skehel JJ, Wiley DC. Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. Nature. 1998;396(6706):92-6.

10. Hellebo A, Vilas U, Falk K, Vlasak R. Infectious salmon anemia virus specifically binds to and hydrolyzes 4-O-acetylated sialic acids. Journal of Virology. 2004;78(6):3055-62.

11. Vlasak R, Krystal M, Nacht M, Palese P. The Influenza-C Virus Glycoprotein (He) Exhibits Receptor-Binding (Hemagglutinin) and Receptor-Destroying (Esterase) Activities. Virology. 1987;160(2):419-25.

12. Herrler G, Durkop I, Becht H, Klenk HD. The glycoprotein of influenza C virus is the haemagglutinin, esterase and fusion factor. The Journal of General Virology. 1988;69 (Pt 4):839-46.

13. Herrler G, Klenk HD. Structure and function of the HEF glycoprotein of influenza C virus. Advances in Virus Research. 1991;40:213-34.

14. Paragas J, Talon J, O'Neill RE, Anderson DK, Garcia-Sastre A, Palese P. Influenza B and C virus NEP (NS2) proteins possess nuclear export activities. Journal of Virology. 2001;75(16):7375-83.

15. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. Evolution and ecology of influenza A viruses. Microbiological Reviews. 1992;56(1):152-79.

16. Hongo S, Sugawara K, Muraki Y, Kitame F, Nakamura K. Characterization of a second protein (CM2) encoded by RNA segment 6 of influenza C virus. Journal of Virology. 1997;71(4):2786-92.

17. Hongo S, Ishii K, Mori K, Takashita E, Muraki Y, Matsuzaki Y, Sugawara K. Detection of ion channel activity in Xenopus laevis oocytes expressing Influenza C virus CM2 protein. Archives of Virology. 2004;149(1):35-50.

18. Muraki Y, Furukawa T, Kohno Y, Matsuzaki Y, Takashita E, Sugawara K, Hongo S. Influenza C virus NS1 protein upregulates the splicing of viral mRNAs. Journal of Virology. 2010;84(4):1957-66.

19. Kohno Y, Muraki Y, Matsuzaki Y, Takashita E, Sugawara K, Hongo S. Intracellular localization of influenza C virus NS2 protein (NEP) in infected cells and its incorporation into virions. Archives of Virology. 2009;154(2):235-43.

20. Osterhaus AD, Rimmelzwaan GF, Martina BE, Bestebroer TM, Fouchier RA. Influenza B virus in seals. Science. 2000;288(5468):1051-3. Epub 2000/05/12.

21. Bodewes R, Morick D, de Mutsert G, Osinga N, Bestebroer T, van der Vliet S, Smits SL, Kuiken T, Rimmelzwaan GF, Fouchier RA, Osterhaus AD. Recurring influenza B virus infections in seals. Emerging Infectious Diseases. 2013;19(3):511-2. Epub 2013/06/12.

22. Zambon MC. Epidemiology and pathogenesis of influenza. J Antimicrobial Chemotherapy. 1999;44 Suppl B:3-9.

23. Calvo C, Garcia-Garcia ML, Centeno M, Perez-Brena P, Casas I. Influenza C virus infection in children, Spain. Emerging Infectious Diseases. 2006;12(10):1621-2.

24. Anton A, Marcos MA, Codoner FM, de Molina P, Martinez A, Cardenosa N, Godoy P, Torner N, Martinez MJ, Ramon S, Tudo G, Isanta R, Gonzalo V, de Anta MT, Pumarola T. Influenza C virus surveillance during the first influenza A (H1N1) 2009 pandemic wave in Catalonia, Spain. Diagnostic Microbiology and Infectious Disease. 2011;69(4):419-27.

25. Tanaka S, Aoki Y, Matoba Y, Yahagi K, Mizuta K, Itagaki T, Katsushima F, Katsushima Y, Matsuzaki Y. The dominant antigenic group of influenza C infections changed from c/Sao Paulo/378/82-lineage to c/Kanagawa/1/76-lineage in Yamagata, Japan, in 2014. Japanese Journal of Infectious Diseases. 2015;68(2):166-8.

26. Matsuzaki Y, Sugawara K, Abiko C, Ikeda T, Aoki Y, Mizuta K, Katsushima N, Katsushima F, Katsushima Y, Itagaki T, Shimotai Y, Hongo S, Muraki Y, Nishimura H. Epidemiological information regarding the periodic epidemics of influenza C virus in Japan (1996-2013) and the seroprevalence of antibodies to different antigenic groups. Journal of Clinical Virology. 2014;61(1):87-93.

27. Kauppila J, Ronkko E, Juvonen R, Saukkoriipi A, Saikku P, Bloigu A, Vainio O, Ziegler T. Influenza C virus infection in military recruits--symptoms and clinical manifestation. Journal of Medical Virology. 2014;86(5):879-85.

28. Salez N, Melade J, Pascalis H, Aherfi S, Dellagi K, Charrel RN, Carrat F, de Lamballerie X. Influenza C virus high seroprevalence rates observed in 3 different population groups. The Journal of Infection. 2014;69(2):182-9.

29. Matsuzaki Y, Abiko C, Mizuta K, Sugawara K, Takashita E, Muraki Y, Suzuki H, Mikawa M, Shimada S, Sato K, Kuzuya M, Takao S, Wakatsuki K, Itagaki T, Hongo S, Nishimura H. A nationwide epidemic of influenza C virus infection in Japan in 2004. Journal of Clinical Microbiology. 2007;45(3):783-8.

30. Matsuzaki Y, Sugawara K, Mizuta K, Tsuchiya E, Muraki Y, Hongo S, Suzuki H, Nakamura K. Antigenic and genetic characterization of influenza C viruses which caused two outbreaks in Yamagata City, Japan, in 1996 and 1998. Journal of Clinical Microbiology. 2002;40(2):422-9.

31. Matsuzaki Y, Mizuta K, Sugawara K, Tsuchiya E, Muraki Y, Hongo S, Suzuki H, Nishimura H. Frequent reassortment among influenza C viruses. Journal of Virology. 2003;77(2):871-81.

32. Odagiri T, Matsuzaki Y, Okamoto M, Suzuki A, Saito M, Tamaki R, Lupisan SP, Sombrero LT, Hongo S, Oshitani H. Isolation and characterization of influenza C viruses in the Philippines and Japan. Journal of Clinical Microbiology. 2014.

33. Pabbaraju K, Wong S, Wong A, May-Hadford J, Tellier R, Fonseca K. Detection of influenza C virus by a real-time RT-PCR assay. Influenza and Other Respiratory Viruses. 2013;7(6):954-60.

34. Gouarin S, Vabret A, Dina J, Petitjean J, Brouard J, Cuvillon-Nimal D, Freymuth F. Study of influenza C virus infection in France. Journal of Medical Virology. 2008;80(8):1441-6.

35. Akinloye OM, Ronkko E, Savolainen-Kopra C, Ziegler T, Iwalokun BA, Deji-Agboola MA, Oluwadun A, Roivainen M, Adu FD, Hovi T. Specific viruses detected in nigerian children in association with acute respiratory disease. Journal of Tropical Medicine. 2011;2011:690286.

36. Matsuzaki Y, Katsushima N, Nagai Y, Shoji M, Itagaki T, Sakamoto M, Kitaoka S, Mizuta K, Nishimura H. Clinical features of influenza C virus infection in children. The Journal of Infectious Diseases. 2006;193(9):1229-35.

37. Joosting AC, Head B, Bynoe ML, Tyrrell DA. Production of common colds in human volunteers by influenza C virus. British Medical Journal. 1968;4(5624):153-4.

38. Keech M, Beardsworth P. The impact of influenza on working days lost: a review of the literature. PharmacoEconomics. 2008;26(11):911-24.

39. Greenbaum E, Morag A, Zakay-Rones Z. Isolation of influenza C virus during an outbreak of influenza A and B viruses. Journal of Clinical Microbiology. 1998;36(5):1441-2.

40. Hirsila M, Kauppila J, Tuomaala K, Grekula B, Puhakka T, Ruuskanen O, Ziegler T. Detection by reverse transcription-polymerase chain reaction of influenza C in nasopharyngeal secretions of adults with a common cold. The Journal of Infectious Diseases. 2001;183(8):1269-72.

41. Moriuchi H, Katsushima N, Nishimura H, Nakamura K, Numazaki Y. Community-acquired influenza C virus infection in children. The Journal of Pediatrics. 1991;118(2):235-8.

42. Katagiri S, Ohizumi A, Ohyama S, Homma M. Follow-up study of type C influenza outbreak in a children's home. Microbiology and Immunology. 1987;31(4):337-43.

43. Calvo C, Garcia-Garcia ML, Borrell B, Pozo F, Casas I. Prospective study of influenza C in hospitalized children. The Pediatric Infectious Disease Journal. 2013;32(8):916-9.

44. Yano T, Maeda C, Akachi S, Matsuno Y, Yamadera M, Kobayashi T, Nagai Y, Iwade Y, Kusuhara H, Katayama M, Fukuta M, Nakagawa Y, Naraya S, Takahashi H, Hiraoka M, Akinori Yamauchi TN, Hideomi Amano, Tetsuo Yamaguchi, Hitoshi Ochiai, Toshiaki Ihara, and Yoko Matsuzaki. Phylogenetic Analysis and Seroprevalence of Influenza C Virus in Mie Prefecture, Japan in 2012. Japanese Journal of Infectious Diseases 2014;67:127-31.

45. Principi N, Scala A, Daleno C, Esposito S. Influenza C virus-associated community-acquired pneumonia in children. Influenza and Other Respiratory Viruses. 2013;7(6):999-1003.

46. Takayanagi M, Umehara N, Watanabe H, Kitamura T, Ohtake M, Nishimura H, Matsuzaki Y, Ichiyama T. Acute encephalopathy associated with influenza C virus infection. The Pediatric Infectious Disease Journal. 2009;28(6):554.

47. Mukherjee RT, Chanda S, Mullick S, De P, Dey-Sarkar M, Chawla-Sarkar M. Spectrum of respiratory viruses circulating in eastern India: prospective surveillance among patients with influenza-like illness during 2010-2011. Journal of Medical Virology. 2013;85(8):1459-65.

48. Hornsleth A, Siggaard-Andersen J, Hjort L. Epidemiology of herpesvirus and respiratory virus infections. Part 1. Serologic findings. Geriatrics. 1975;30(8):61-8.

49. Homma M, Ohyama S, Katagiri S. Age distribution of the antibody to type C influenza virus. Microbiology and Immunology. 1982;26(7):639-42.

50. Shu YL, Li CKF, Li Z, Gao RB, Liang Q, Zhang Y, Dong LB, Zhou JF, Dong J, Wang DY, Wen LY, Wang M, Bai T, Li DX, Dong XP, Yu HJ, Yang WZ, Wang Y, Feng ZJ, McMichael AJ, Xu XN. Avian Influenza A(H5N1) Viruses Can Directly Infect and Replicate in Human Gut Tissues. Journal of Infectious Diseases. 2010;201(8):1173-7.

51. Campbell NA. Biology: a Global Approach. Tenth edition / Neil A. Campbell [Global edition. Ed.

52. Human influenza: aspects of the immune response to vaccination. Annals of Internal Medicine. 1969;71(2):369-98.

53. Jolliff CR, Cost KM, Stivrins PC, Grossman PP, Nolte CR, Franco SM, Fijan KJ, Fletcher LL, Shriner HC. Reference Intervals for Serum Igg, Iga, Igm, C-3, and C-4 as Determined by Rate Nephelometry. Clinical Chemistry. 1982;28(1):126-8.

54. Agarwal S, Cunningham-Rundles C. Assessment and clinical interpretation of reduced IgG values. Annals of Allergy, Asthma & Immunology. 2007;99(3):281-3.

55. Crum-Cianflone NF, Collins G, Defang G, Iverson E, Eberly LE, Duplessis C, Maguire J, Ganesan A, Agan BK, Lalani T, Whitman T, Brandt C, Faix D, Blair PJ, Burgess T. Immunoglobulin G subclass levels and antibody responses to the 2009 influenza A (H1N1) monovalent vaccine among human immunodeficiency virus (HIV)-infected and HIV-uninfected adults. Clinical and Experimental Immunology. 2012;168(1):135-41.

56. Puissant-Lubrano B, Peres M, Apoil PA, Congy-Jolivet N, Roubinet F, Blancher A. Immunoglobulin IgA, IgD, IgG, IgM and IgG subclass reference values in adults. Clinical Chemistry and Laboratory Medicine. 2015.

57. Jegaskanda S, Reading PC, Kent SJ. Influenza-specific antibody-dependent cellular cytotoxicity: toward a universal influenza vaccine. Journal of Immunology. 2014;193(2):469-75.

58. Gatherer D. Tempo and mode in the molecular evolution of influenza C. PLoS Currents. 2010;2:RRN1199.

59. Peng G, Hongo S, Muraki Y, Sugawara K, Nishimura H, Kitame F, Nakamura K. Genetic reassortment of influenza C viruses in man. The Journal of General Virology. 1994;75 (Pt 12):3619-22.

60. Peng G, Hongo S, Kimura H, Muraki Y, Sugawara K, Kitame F, Numazaki Y, Suzuki H, Nakamura K. Frequent occurrence of genetic reassortment between influenza C virus strains in nature. Journal of General Virology. 1996;77:1489-92.

61. Matsuzaki Y, Mizuta K, Kimura H, Sugawara K, Tsuchiya E, Suzuki H, Hongo S, Nakamura K. Characterization of antigenically unique influenza C virus strains isolated in

Yamagata and Sendai cities, Japan, during 1992-1993. The Journal of General Virology. 2000;81(Pt 6):1447-52.

62. Tada Y, Hongo S, Muraki Y, Sugawara K, Kitame F, Nakamura K. Evolutionary analysis of influenza C virus M genes. Virus Genes. 1997;15(1):53-9.

63. Alamgir AS, Matsuzaki Y, Hongo S, Tsuchiya E, Sugawara K, Muraki Y, Nakamura K. Phylogenetic analysis of influenza C virus nonstructural (NS) protein genes and identification of the NS2 protein. The Journal of General Virology. 2000;81(Pt 8):1933-40.

64. Chen R, Holmes EC. The evolutionary dynamics of human influenza B virus. Journal of Molecular Evolution. 2008;66(6):655-63.

65. Matsuzaki Y, Takao S, Shimada S, Mizuta K, Sugawara K, Takashita E, Muraki Y, Hongo S, Nishimura H. Characterization of antigenically and genetically similar influenza C viruses isolated in Japan during the 1999-2000 season. Epidemiology and Infection. 2004;132(4):709-20.

66. Treanor J. Influenza vaccine--outmaneuvering antigenic shift and drift. New England Journal of Medicine. 2004;350(3):218-20.

67. Fernando C Motta, Luiz MO, Couceiro JN. Serological analysis reveals circulation of influenza C viruses, Brazil. Journal of Public Health. 2000;34(2):204-5.

68. R. J. O'Callaghani, D. D. Labat. Human Antibody to Influenza C Virus: Its Age-Related Distribution and Distinction from Receptor Analogs. Infection and Immunity. 1980;30(2):500-5.

69. Kaji M, Hiromatsu Y, Kashiwagi S, Hayashi J, Oyama S, Katagiri S, Homma M. Distribution of Antibodies to Influenza C Virus. The Kurume Medical Journal. 1983;30:121-3.

70. Andrews BE, McDonald JC. Influenza virus C infection in England. British Medical Journal. 1955;2(4946):992-4.

71. Katagiri S, Ohizumi A, Homma M. An outbreak of type C influenza in a children's home. The Journal of Infectious Diseases. 1983;148(1):51-6.

72. M. R. Hilleman JHW, R.L. Gauld. Influenza Antibodies In The Population Of The USA An Epidemiological Investigation. Bulletin of World Health Organisation. 1953:613-31.

73. Buonagurio DA, Nakada S, Fitch WM, Palese P. Epidemiology of influenza C virus in man: multiple evolutionary lineages and low rate of change. Virology. 1986;153(1):12-21.

74. Mukherjee RT, Mukherjee A, Mullick S, Chawla-Sarkar M. Full genome analysis and characterization of influenza C virus identified in Eastern India. Infection, Genetics and Evolution. 2013;16:419-25.

75. Vejtorp M, Leerhoy J. Rubella IgG antibody detection by ELISA using capillary blood samples collected on filter paper and in microtainer tubes. Acta Pathologica et Microbiologica Scandinavica Section B, Microbiology. 1981;89(5):369-70.

76. Vaheri A, Vaananen P, Salonen EM, Suni J. Rubella antibody determination from heparinised finger-tip blood by single radial haemolysis and enzyme immunoassay. Journal of Clinical Pathology. 1980;33(9):845-7.

77. Voller A, Bidwell DE. A simple method for detecting antibodies to rubella. British Journal of Experimental Pathology. 1975;56(4):338-9.

78. Onyango CO, Njeru R, Kazungu S, Achilla R, Bulimo W, Welch SR, Cane PA, Gunson RN, Hammitt LL, Scott JA, Berkley JA, Nokes DJ. Influenza surveillance among children with pneumonia admitted to a district hospital in coastal Kenya, 2007-2010. The Journal of Infectious Diseases. 2012;206 Suppl 1:S61-7.

79. Covalciuc KA, Webb KH, Carlson CA. Comparison of four clinical specimen types for detection of influenza A and B viruses by optical immunoassay (FLU OIA test) and cell culture methods. Journal of Clinical Microbiology. 1999;37(12):3971-4.

80. Jeong JH, Kim KH, Jeong SH, Park JW, Lee SM, Seo YH. Comparison of sputum and nasopharyngeal swabs for detection of respiratory viruses. Journal of Medical Virology. 2014;86(12):2122-7.

81. Baron S, Peake RC, Schuenke S. Medical Microbiology. 4th ed / associate editor Rhonda C. Peake / scientific illustrator Steve Schuenke. ed. Galveston, Tex.: University of Texas Medical Branch at Galveston; 1996.

82. Wertheim HF, Melles DC, Vos MC, van Leeuwen W, van Belkum A, Verbrugh HA, Nouwen JL. The role of nasal carriage in Staphylococcus aureus infections. The Lancet Infectious Diseases. 2005;5(12):751-62.

83. Kluytmans J, van Belkum A, Verbrugh H. Nasal carriage of Staphylococcus aureus: epidemiology, underlying mechanisms, and associated risks. Clinical Microbiology Reviews. 1997;10(3):505-20.

84. Eriksen NHR, Espersen F, Rosdahl VT, Jensen K. Carriage of Staphylococcus-Aureus among 104 Healthy-Persons during a 19-Month Period. Epidemiology and Infection. 1995;115(1):51-60.

85. Herwaldt LA, Cullen JJ, French P, Hu JF, Pfaller MA, Wenzel RP, Perl TM. Preoperative risk factors for nasal carriage of Staphylococcus aureus. Infection Control and Hospital Epidemiology. 2004;25(6):481-4.

86. Peacock SJ, Justice A, Griffiths D, de Silva GDI, Kantzanou MN, Crook D, Sleeman K, Day NPJ. Determinants of acquisition and carriage of Staphylococcus aureus in infancy. Journal of Clinical Microbiology. 2003;41(12):5718-25.

87. Armstrongesther CA, Smith JE. Carriage Patterns of Staphylococcus-Aureus in a Healthy Non-Hospital Population of Adults and Children. Annals of Human Biology. 1976;3(3):221-7.

88. Solberg CO. A Study of Carriers of Staphylococcus Aureus - with Special Regard to Quantitative Bacterial Estimations. Acta Medica Scandinavica. 1965;S178:9-+.

89. Wertheim HFL, Van Kleef M, Vos MC, Ott A, Verbrugh HA, Fokkens W. Nose picking and nasal carriage of Staphylococcus aureus. Infection Control and Hospital Epidemiology. 2006;27(8):863-7.

90. Reuter JA, Spacek DV, Snyder MP. High-throughput sequencing technologies. Molecular Cell. 2015;58(4):586-97. Epub 2015/05/23.

91. Phan TG, Luchsinger V, Avendano LF, Deng X, Delwart E. Cyclovirus in nasopharyngeal aspirates of Chilean children with respiratory infections. The Journal of General Virology. 2014;95(Pt 4):922-7. Epub 2014/01/15.

92. Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, Ostell J, Lipman D. The influenza virus resource at the National Center for Biotechnology Information. Journal of Virology. 2008;82(2):596-601.

93. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990;215(3):403-10. Epub 1990/10/05.

94. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. Artemis: sequence visualization and annotation. Bioinformatics. 2000;16(10):944-5. Epub 2000/12/20.

95. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution. 2012;29(8):1969-73. Epub 2012/03/01.

96. Rambaut A. Path-O-Gen v1.4 2015 [cited 2015 28/7/15]; Available from: http://tree.bio.ed.ac.uk/software/pathogen/.

97. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution. 2013;30(12):2725-9. Epub 2013/10/18.

98. Rambaut AS, MA. Xie, D. Drummond, AJ. Tracer v1.6. 2014.

99. Rambaut A. FigTree v1.4. 2015 [cited 2015 23/7/15]; Available from: http://tree.bio.ed.ac.uk/software/figtree/.

100. Wheeler DL, Church DM, Edgar R, Federhen S, Helmberg W, Madden TL, Pontius JU, Schuler GD, Schriml LM, Sequeira E, Suzek TO, Tatusova TA, Wagner L. Database resources of the National Center for Biotechnology Information: update. Nucleic Acids Research. 2004;32(Database issue):D35-40. Epub 2003/12/19.

101. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Research. 2011;39(Database issue):D32-7. Epub 2010/11/13.

102. Drummond AJ, Rambaut A, Shapiro B, Pybus OG. Bayesian coalescent inference of past population dynamics from molecular sequences. Molecular Biology and Evolution. 2005;22(5):1185-92. Epub 2005/02/11.

103. Salauze D, Serre V, Perrin C. Quantification of Total Igm and Igg Levels in Rat Sera by a Sandwich Elisa Technique. Comparative Haematology International. 1994;4(1):30-3.

104. Nakayama E, Yokoyama A, Miyamoto H, Igarashi M, Kishida N, Matsuno K, Marzi A, Feldmann H, Ito K, Saijo M, Takada A. Enzyme-linked immunosorbent assay for detection of filovirus species-specific antibodies. Clinical and Vaccine Immunology. 2010;17(11):1723-8.

105. Shetty V, Trumbull K, Hegde A, Shenoy V, Prabhu R, K S, Palavecino E, Shetty AK. Prevalence of Community-Acquired Methicillin-Resistant Staphylococcus aureus Nasal Colonization Among Children. Journal of Clinical and Diagnostic Research. 2014;8(12):DC12-5.

106. Saxena S, Singh K, Talwar V. Methicillin-resistant Staphylococcus aureus prevalence in community in the east Delhi area. Japanese Journal of Infectious Diseases. 2003;56(2):54-6.

107. Pathak A, Marothi Y, Iyer RV, Singh B, Sharma M, Eriksson B, Macaden R, Lundborg CS. Nasal carriage and antimicrobial susceptibility of Staphylococcus aureus in healthy preschool children in Ujjain, India. BMC Pediatrics. 2010;10:100.

108. Ramana KV, Mohanty SK, Wilson CG. Staphylococcus aureus colonization of anterior nares of school going children. Indian Journal of Pediatrics. 2009;76(8):813-6.

109. Dey S, Rosales-Klintz S, Shouche S, Pathak JP, Pathak A. Prevalence and risk factors for nasal carriage of Staphylococcus aureus in children attending anganwaries (preschools) in Ujjain, India. BMC Research Notes. 2013;6:265.

110. Chatterjee SS, Ray P, Aggarwal A, Das A, Sharma M. A community-based study on nasal carriage of Staphylococcus aureus. The Indian Journal of Medical Research. 2009;130(6):742-8.

Appendices

Appendix 1

Samples sent for deep sequencing

A - Paediatric and symptomatic (lower RNA reading on NanoDrop)

- K012
- K074
- K134

B - Paediatric and symptomatic (higher RNA reading on NanoDrop)

- K014
- K016

C - Adult, high fluC IgG, symptomatic

- K008
- K032
- K036
- K041
- K043
- K044
- K045
- K049
- K137
- K145

D - Adult and fever

- K025
- K033
- K065
- K097
- K112
- K132
- K141
- K146

E – Asymptomatic participants, low total IgG

- K004
- K005
- K006
- K007

- K009
- K020
- K066
- K072
- K118
- K131

F - Asthmatics (higher fluC IgG)

- K021
- K031
- K042
- K048
- K050
- K067
- K080
- K083
- K140
- K143

G - COPD

- K017
- K090
- K094
- K105
- K139
- K147

H - Positive PCR

- K081

I - Positive PCR

- K087

Appendix 2

Table 6 – Influenza C IgG Concentration

This table shows the mean fluC IgG concentration in mg/dl for each participant who provided a serum sample. Samples which were classed as positive were coloured green and those which were classed as negative were coloured red. The 'estimation above' column is for those samples which appeared above the IgG curve, when plotted, and so the concentration of the IgG samples cannot be accurately measured above the figure given. A dash (-) indicates that a serum sample was not provided.

Sample no.	IgG concentration (mg/dl)	Estimation above:	Positive/Negative
K001	0.94		Negative
K002	4.01		Positive
K003	0.98		Negative
K004	3.07		Positive
K005	1.36		Negative
K006	2.42		Positive
K007	2.75		Positive
K008	5.06		Positive
K009	3.58		Positive
K010	1.94		Positive
K011	3.04		Positive
K012	-		
K013	3.38		Positive
K014	-		
K015	3.54		Positive
K016	-		
K017	7.47	5.88	Positive
K018	1.86		Positive
K019	3.13		Positive

K020	5.37		Positive
K021	4.72		Positive
K022	-		
K023	-		
K024	-		
K025	-		
K026	0.56		Negative
K027	4.37		Positive
K028	2.94		Positive
K029	3.56		Positive
K030	5.97	5.88	Positive
K031	7.51	5.88	Positive
K032	5.77	3.12	Positive
K033	-		
K034	1.35		Positive
K035	5.96	3.12	Positive
K036	4.39	3.12	Positive
K037	3.81	3.12	Positive
K038	1.80		Positive
K039	3.50	3.12	Positive
K040	1.39		Positive
K041	4.27	3.12	Positive
K042	2.83		Positive
K043	5.28	3.12	Positive
K044	4.57	3.12	Positive
K045	4.98	3.12	Positive
K046	-		
K047	4.09	3.12	Positive
K048	3.25	3.12	Positive
K049	4.31	3.12	Positive
K050	1.93		Positive

K051	5.18	3.12	Positive
K052	3.59	3.12	Positive
K053	1.71		Positive
K054	3.60	3.12	Positive
K055	2.15		Positive
K056	1.50		Positive
K057	0.74		Negative
K058	3.27	3.12	Positive
K059	1.59		Positive
K060	1.51		Positive
K061	2.91		Positive
K062	3.24	3.12	Positive
K063	2.49		Positive
K064	2.83		Positive
K065	2.67		Positive
K066	1.60		Positive
K067	2.49		Positive
K068	0.74		Negative
K069	1.61		Positive
K070	0.28		Negative
K071	0.77		Negative
K072	0.39		Negative
K073	1.57		Positive
K074	-		
K075	0.67		Negative
K076	0.89		Negative
K077	3.04		Positive
K078	2.76		Positive
K079	2.49		Positive
K080	2.90		Positive
K081	2.99		Positive

K082	2.62		Positive
K083	2.59		Positive
K084	-		
K085	3.24	3.12	Positive
K086	1.16		Positive
K087	1.68		Positive
K088	0.83		Positive
K089	2.54		Positive
K090	2.20		Positive
K091	0.84		Positive
K092	3.58	3.12	Positive
K093	0.77		Negative
K094	2.96		Positive
K095	0.24		Negative
K096	1.42		Positive
K097	1.41		Positive
K098	2.21		Positive
K099	2.58		Positive
K100	2.33		Positive
K101	0.62		Negative
K102	-		
K103	-		
K104	-		
K105	-		
K106	2.13		Positive
K107	-		
K108	2.05		Positive
K109	1.12		Positive
K110	3.39	3.12	Positive
K111	3.69	3.12	Positive
K112	3.27	3.12	Positive

K113	2.29		Positive
K114	2.98		Positive
K115	4.10	3.13	Positive
K116	1.14		Negative
K117	1.08		Negative
K118	0.57		Negative
K119	2.03		Positive
K120	3.19	3.13	Positive
K121	0.28		Negative
K122	2.35		Positive
K123	1.41		Positive
K124	2.16		Positive
K125	2.57		Positive
K126	2.73		Positive
K127	1.77		Positive
K128	0.46		Negative
K129	2.53		Positive
K130	1.51		Positive
K131	1.69		Positive
K132	2.79		Positive
K133	0.78		Negative
K134	-		
K135	0.94		Negative
K136	4.19	3.13	Positive
K137	5.17	3.13	Positive
K138	2.86		Positive
K139	2.70		Positive
K140	4.48	3.12	Positive
K141	-		
K142	4.54	3.12	Positive
K143	5.35	3.12	Positive

K144	1.77		Negative
K145	5.89	3.12	Positive
K146	-		
K147	1.90		Negative
K148	4.54	3.12	Positive

Appendix 3

Table 10 – Total IgG antibody results

This table shows the mean total IgG concentration in mg/dl for each participant who provided a serum sample. Samples which were classed as below the normal range were coloured green and those which were classed as above were coloured red. A dash (-) indicates that a serum sample was not provided.

Sample no.	IgG concentration (mg/dl)
K001	2909.66
K002	695.24
K003	4924.72
K004	595.18
K005	637.64
K006	293.39
K007	483.41
K008	658.41
K009	612.60
K010	732.69
K011	1007.91
K012	-
K013	577.23
K014	-
K015	854.46
K016	-
K017	431.68
K018	424.14
K019	633.59
K020	318.91
K021	413.00

K022	-
K023	-
K024	-
K025	-
K026	462.91
K027	409.12
K028	379.68
K029	782.83
K030	2834.05
K031	2997.30
K032	2302.18
K033	-
K034	1931.07
K035	2341.04
K036	1516.13
K037	3289.27
K038	1727.36
K039	3246.99
K040	767.83
K041	2502.14
K042	3486.47
K043	4364.84
K044	3811.76
K045	2820.84
K046	-
K047	2258.79
K048	2040.67
K049	2161.52
K050	3147.10
K051	3382.18
K052	2054.70

K053	3173.90
K054	4751.15
K055	3881.72
K056	3617.08
K057	2838.12
K058	3753.35
K059	2242.22
K060	5182.95
K061	3205.10
K062	3433.38
K063	905.45
K064	670.33
K065	1308.37
K066	449.46
K067	2661.03
K068	676.09
K069	1024.45
K070	939.67
K071	1150.95
K072	572.16
K073	2998.30
K074	-
K075	3673.23
K076	2716.19
K077	3805.44
K078	2862.63
K079	5400.34
K080	5504.65
K081	4198.77
K082	3213.94
K083	4481.09

K084	-
K085	3664.54
K086	2723.17
K087	2701.20
K088	3264.07
K089	3580.50
K090	2678.07
K091	3877.32
K092	3134.89
K093	4607.99
K094	3034.23
K095	1194.62
K096	1644.44
K097	1589.74
K098	1284.18
K099	2838.31
K100	2180.33
K101	3264.56
K102	-
K103	-
K104	-
K105	-
K106	1157.08
K107	-
K108	1337.17
K109	1041.22
K110	1157.36
K111	1780.58
K112	756.74
K113	1141.43
K114	1527.09

K115	1966.00
K116	865.59
K117	1165.90
K118	550.04
K119	657.90
K120	2722.52
K121	3793.68
K122	2663.21
K123	876.70
K124	1524.80
K125	1447.43
K126	1144.51
K127	2464.01
K128	978.06
K129	854.20
K130	631.72
K131	642.75
K132	1024.84
K133	624.76
K134	-
K135	745.72
K136	1240.99
K137	591.61
K138	629.62
K139	486.98
K140	1720.23
K141	-
K142	1695.54
K143	1207.33
K144	747.83
K145	1645.34

K146	-
K147	2122.29
K148	1203.35

Appendix 4

Table 11 – Comparison of FluC and Total IgG Antibodies

This table compares the mean fluC IgG concentration with the mean total IgG concentration, as a percentage. A dash (-) indicates that a serum sample was not provided.

Sample no.	IgG concentration (mg/dl)	Estimation above:	IgG concentration (mg/dL)	%
K001	0.94		2909.66	0.03
K002	4.01		695.24	0.58
K003	0.98		4924.72	0.02
K004	3.07		595.18	0.52
K005	1.36		637.64	0.21
K006	2.42		293.39	0.82
K007	2.75		483.41	0.57
K008	5.06		658.41	0.77
K009	3.58		612.60	0.58
K010	1.94		732.69	0.26
K011	3.04		1007.91	0.30
K012	-		-	-
K013	3.38		577.23	0.58
K014	-		-	-
K015	3.54		854.46	0.41
K016	-		-	-
K017	7.47	5.8773437	431.68	1.36
K018	1.86		424.14	0.44
K019	3.13		633.59	0.49
K020	5.37		318.91	1.68
K021	4.72		413.00	1.14
K022	-		-	-
K023	-		-	-

K024	-		-	-
K025	-		-	-
K026	0.56		462.91	0.12
K027	4.37		409.12	1.07
K028	2.94		379.68	0.78
K029	3.56		782.83	0.46
K030	5.97	5.8773437	2834.05	0.21
K031	7.51	5.8773437	2997.30	0.20
K032	5.77	3.1229508	2302.18	0.14
K033	-		-	-
K034	1.35		1931.07	0.07
K035	5.96	3.1229508	2341.04	0.13
K036	4.39	3.1229508	1516.13	0.21
K037	3.81	3.1229508	3289.27	0.09
K038	1.80		1727.36	0.10
K039	3.50	3.1229508	3246.99	0.10
K040	1.39		767.83	0.18
K041	4.27	3.1229508	2502.14	0.12
K042	2.83		3486.47	0.08
K043	5.28	3.1229508	4364.84	0.07
K044	4.57	3.1229508	3811.76	0.08
K045	4.98	3.1229508	2820.84	0.11
K046	-		-	-
K047	4.09	3.1229508	2258.79	0.14
K048	3.25	3.1229508	2040.67	0.15
K049	4.31	3.1229508	2161.52	0.14
K050	1.93		3147.10	0.06
K051	5.18	3.1229508	3382.18	0.09
K052	3.59	3.1230115	2054.70	0.15
K053	1.71		3173.90	0.05
K054	3.60	3.1230115	4751.15	0.07

K055	2.15		3881.72	0.06
K056	1.50		3617.08	0.04
K057	0.74		2838.12	0.03
K058	3.27	3.1230115	3753.35	0.08
K059	1.59		2242.22	0.07
K060	1.51		5182.95	0.03
K061	2.91		3205.10	0.09
K062	3.24	3.1230115	3433.38	0.09
K063	2.49		905.45	0.28
K064	2.83		670.33	0.42
K065	2.67		1308.37	0.20
K066	1.60		449.46	0.36
K067	2.49		2661.03	0.09
K068	0.74		676.09	0.11
K069	1.61		1024.45	0.16
K070	0.28		939.67	0.03
K071	0.77		1150.95	0.07
K072	0.39		572.16	0.07
K073	1.57		2998.30	0.05
K074	-		-	-
K075	0.67		3673.23	0.02
K076	0.89		2716.19	0.03
K077	3.04		3805.44	0.08
K078	2.76		2862.63	0.10
K079	2.49		5400.34	0.05
K080	2.90		5504.65	0.05
K081	2.99		4198.77	0.07
K082	2.62		3213.94	0.08
K083	2.59		4481.09	0.06
K084	-		-	-
K085	3.24	3.1229741	3664.54	0.09

K086	1.16		2723.17	0.04
K087	1.68		2701.20	0.06
K088	0.83		3264.07	0.03
K089	2.54		3580.50	0.07
K090	2.20		2678.07	0.08
K091	0.84		3877.32	0.02
K092	3.58	3.1229741	3134.89	0.10
K093	0.77		4607.99	0.02
K094	2.96		3034.23	0.10
K095	0.24		1194.62	0.02
K096	1.42		1644.44	0.09
K097	1.41		1589.74	0.09
K098	2.21		1284.18	0.17
K099	2.58		2838.31	0.09
K100	2.33		2180.33	0.11
K101	0.62		3264.56	0.02
K102	-		-	-
K103	-		-	-
K104	-		-	-
K105	-		-	-
K106	2.13		1157.08	0.18
K107	-		-	-
K108	2.05		1337.17	0.15
K109	1.12		1041.22	0.11
K110	3.39	3.1230115	1157.36	0.27
K111	3.69	3.1230115	1780.58	0.18
K112	3.27	3.1230115	756.74	0.41
K113	2.29		1141.43	0.20
K114	2.98		1527.09	0.19
K115	4.10	3.127305	1966.00	0.16
K116	1.14		865.59	0.13

K117	1.08		1165.90	0.09
K118	0.57		550.04	0.10
K119	2.03		657.90	0.31
K120	3.19	3.127305	2722.52	0.11
K121	0.28		3793.68	0.01
K122	2.35		2663.21	0.09
K123	1.41		876.70	0.16
K124	2.16		1524.80	0.14
K125	2.57		1447.43	0.18
K126	2.73		1144.51	0.24
K127	1.77		2464.01	0.07
K128	0.46		978.06	0.05
K129	2.53		854.20	0.30
K130	1.51		631.72	0.24
K131	1.69		642.75	0.26
K132	2.79		1024.84	0.27
K133	0.78		624.76	0.12
K134	-		-	-
K135	0.94		745.72	0.13
K136	4.19	3.127305	1240.99	0.25
K137	5.17	3.127305	591.61	0.53
K138	2.86		629.62	0.45
K139	2.70		486.98	0.56
K140	4.48	3.1249885	1720.23	0.18
K141	-		-	-
K142	4.54	3.1249885	1695.54	0.18
K143	5.35	3.1249885	1207.33	0.26
K144	1.77		747.83	0.24
K145	5.89	3.1249885	1645.34	0.19
K146	-		-	-
K147	1.90		2122.29	0.09

K148	4.54	3.1249885	1203.35	0.26

Appendix 5

Table 15 – *Staphylococcus* plate results

This table shows the results of the *Staphylococcus*-selective plates, used to detect the presence of *Staphylococci* on the nasal and nasopharyngeal swabs. A result of 'Y' indicates positive growth and 'N' indicates no growth. N/A indicates that a plate result was not available for that participant.

Sample	Date streaked	Date checked	*Staphylococcus* present?	Coagulase +ve?
K001	N/A	N/A	N/A	N/A
K002	N/A	N/A	N/A	N/A
K003	N/A	N/A	N/A	N/A
K004	16/12/2014	18/12/2014	Y	N
K005	16/12/2014	18/12/2014	Y	Y
K006	17/12/2014	19/12/2014	Y	Y
K007	17/12/2014	19/12/2014	Y	Y
K008	17/12/2014	19/12/2014	Y	Y
K009	17/12/2014	19/12/2014	Y	Y
K010	17/12/2014	19/12/2014	Y	N
K011	17/12/2014	19/12/2014	Y	Y
K012	17/12/2014	19/12/2014	N	N
K013	17/12/2014	19/12/2014	Y	Y
K014	18/12/2014	19-22/12/14	Y	Y
K015	18/12/2014	19-22/12/14	Y	N
K016	18/12/2014	19-22/12/14	Y	N
K017	06/01/2015	08/01/2015	Y	N
K018	06/01/2015	08/01/2015	Y	Y
K019	06/01/2015	08/01/2015	Y	N
K020	06/01/2015	08/01/2015	Y	Y

K021	06/01/2015	08/01/2015	Y	Y
K022	06/01/2015	08/01/2015	Y	Y
K023	06/01/2015	08/01/2015	Y	N
K024	06/01/2015	08/01/2015	Y	Y
K025	06/01/2015	08/01/2015	Y	Y
K026	N/A	N/A	N/A	N/A
K027	N/A	N/A	N/A	N/A
K028	06/01/2015	08/01/2015	Y	Y
K029	08/01/2015	12/01/2015	Y	N
K030	08/01/2015	12/01/2015	Y	Y
K031	08/01/2015	12/01/2015	Y	Y
K032	08/01/2015	12/01/2015	Y	Y
K033	08/01/2015	12/01/2015	N	N
K034	09/01/2015	12/01/2015	Y	Y
K035	13/01/2015	15/01/2015	Y	N
K036	13/01/2015	15/01/2015	N	N
K037	13/01/2015	15/01/2015	Y	Y
K038	N/A	N/A	N/A	N/A
K039	N/A	N/A	N/A	N/A
K040	N/A	N/A	N/A	N/A
K041	13/01/2015	15/01/2015	Y	Y
K042	13/01/2015	14/01/2015	Y	Y
K043	13/01/2015	15/01/2015	Y	Y
K044	13/01/2015	15/01/2015	Y	N
K045	14/01/2015	16/01/2015	Y	N
K046	14/01/2015	15/01/2015	Y	Y
K047	14/01/2015	15/01/2015	Y	Y
K048	14/01/2015	15/01/2015	Y	Y
K049	14/01/2015	15/01/2015	Y	Y
K050	14/01/2015	15/01/2015	Y	Y
K051	15/01/2015	19/01/2015	Y	Y

K052	15/01/2015	19/01/2015	N	N
K053	16/01/2015	19/01/2015	Y	Y
K054	16/01/2015	19/01/2015	Y	Y
K055	16/01/2015	19/01/2015	Y	Y
K056	16/01/2015	19/01/2015	Y	Y
K057	16/01/2015	19/01/2015	Y	Y
K058	16/01/2015	19/01/2015	Y	N
K059	16/01/2015	19/01/2015	Y	Y
K060	16/01/2015	19/01/2015	Y	Y
K061	16/01/2015	19/01/2015	Y	N
K062	16/01/2015	19/01/2015	Y	Y
K063	16/01/2015	19/01/2015	Y	Y
K064	16/01/2015	19/01/2015	Y	N
K065	19/01/2015	21/01/2015	Y	N
K066	19/01/2015	21/01/2015	Y	N
K067	20/01/2015	21/01/2015	Y	Y
K068	20/01/2015	23/01/2015	Y	Y
K069	20/01/2015	23/01/2015	Y	Y
K070	20/01/2015	23/01/2015	Y	Y
K071	20/01/2015	23/01/2015	Y	Y
K072	20/01/2015	21/01/2015	Y	Y
K073	20/01/2015	21/01/2015	Y	Y
K074	20/01/2015	23/01/2015	N	N
K075	21/01/2015	23/01/2015	N	N
K076	21/01/2015	23/01/2015	Y	N
K077	21/01/2015	23/01/2015	Y	Y
K078	22/01/2015	26/01/2015	Y	Y
K079	22/01/2015	26/01/2015	Y	Y
K080	22/01/2015	26/01/2015	Y	Y
K081	22/01/2015	26/01/2015	Y	Y
K082	22/01/2015	26/01/2015	Y	N

K083	22/01/2015	26/01/2015	Y	Y
K084	23/01/2015	26/01/2015	Y	Y
K085	23/01/2015	26/01/2015	Y	Y
K086	23/01/2015	26/01/2015	Y	N
K087	23/01/2015	26/01/2015	Y	Y
K088	26/01/2015	28/01/2015	N	N
K089	26/01/2015	28/01/2015	Y	Y
K090	26/01/2015	28/01/2015	Y	Y
K091	26/01/2015	28/01/2015	Y	Y
K092	26/01/2015	28/01/2015	Y	Y
K093	26/01/2015	28/01/2015	Y	Y
K094	26/01/2015	28/01/2015	Y	Y
K095	26/01/2015	28/01/2015	Y	Y
K096	28/01/2015	30/01/2015	Y	N
K097	28/01/2015	30/01/2015	Y	N
K098	28/01/2015	30/01/2015	Y	N
K099	28/01/2015	30/01/2015	Y	N
K100	28/01/2015	30/01/2015	Y	N
K101	28/01/2015	30/01/2015	Y	Y
K102	28/01/2015	30/01/2015	Y	Y
K103	29/01/2015	02/02/2015	Y	Y
K104	29/01/2015	02/02/2015	Y	N
K105	29/01/2015	02/02/2015	Y	Y
K106	29/01/2015	02/02/2015	Y	Y
K107	29/01/2015	02/02/2015	Y	N
K108	30/01/2015	02/02/2015	Y	N
K109	30/01/2015	02/02/2015	Y	Y
K110	02/02/2015	05/02/2015	Y	N
K111	02/02/2015	04/02/2014	Y	Y
K112	02/02/2015	04/02/2015	Y	Y
K113	02/02/2015	05/02/2015	Y	N

K114	02/02/2015	03/02/2015	Y	Y
K115	02/02/2015	03/02/2015	Y	Y
K116	02/02/2015	05/02/2015	Y	N
K117	04/02/2015	06/02/2015	Y	Y
K118	04/02/2015	06/02/2015	N	N
K119	04/02/2015	06/02/2015	N	N
K120	04/02/2015	06/02/2015	N	N
K121	04/02/2015	06/02/2015	Y	N
K122	04/02/2015	06/02/2015	Y	N
K123	04/02/2015	06/02/2015	Y	Y
K124	05/02/2015	06/02/2015	Y	Y
K125	05/02/2015	09/02/2015	Y	Y
K126	05/02/2015	09/02/2015	N	N
K127	05/02/2015	09/02/2015	Y	N
K128	05/02/2015	09/02/2015	Y	Y
K129	05/02/2015	09/02/2015	Y	Y
K130	05/02/2015	06/02/2015	Y	Y
K131	06/02/2015	09/02/2015	Y	Y
K132	06/02/2015	09/02/2015	Y	Y
K133	06/02/2015	09/02/2015	Y	Y
K134	06/02/2015	09/02/2015	Y	Y
K135	09/02/2015	11/02/2015	Y	N
K136	12/02/2015	16/02/2015	Y	Y
K137	12/02/2015	16/02/2015	Y	Y
K138	12/02/2015	16/02/2015	Y	Y
K139	17/02/2015	19/02/2015	Y	N
K140	17/02/2015	19/02/2015	Y	Y
K141	17/02/2015	19/02/2015	Y	Y
K142	17/02/2015	19/02/2015	Y	N
K143	19/02/2015	23/02/2015	Y	Y
K144	19/02/2015	20/02/2015	Y	Y

K145	23/02/2015	25/02/2015	N	N
K146	23/02/2015	25/02/2015	Y	N
K147	25/02/2015	27/02/2015	Y	Y
K148	05/05/2015	07/05/2015	Y	Y