Appendix for Cohort Profile: The China Aging Respiratory infections Study (CARES): a prospective cohort study in Eastern China

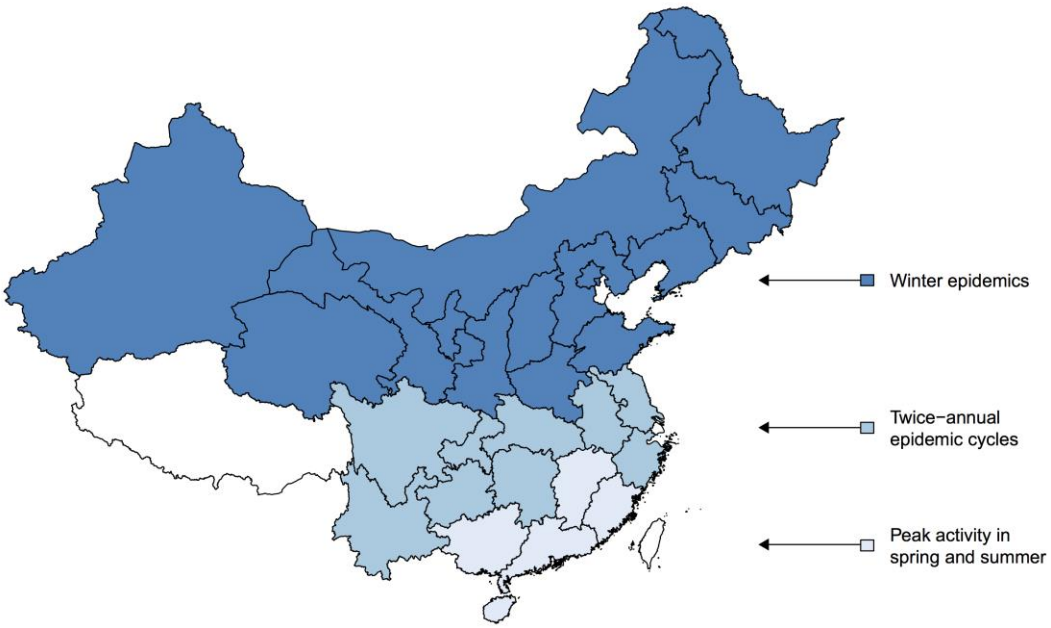
This supplement provides additional details of the approaches to enrolment and follow-up of participants, and laboratory methods used. Data capture forms are included as annexes, at the end of this appendix.

Table of Contents **Page**

1. Selection of study sites

1.1. Influenza Circulation and Province Selection.

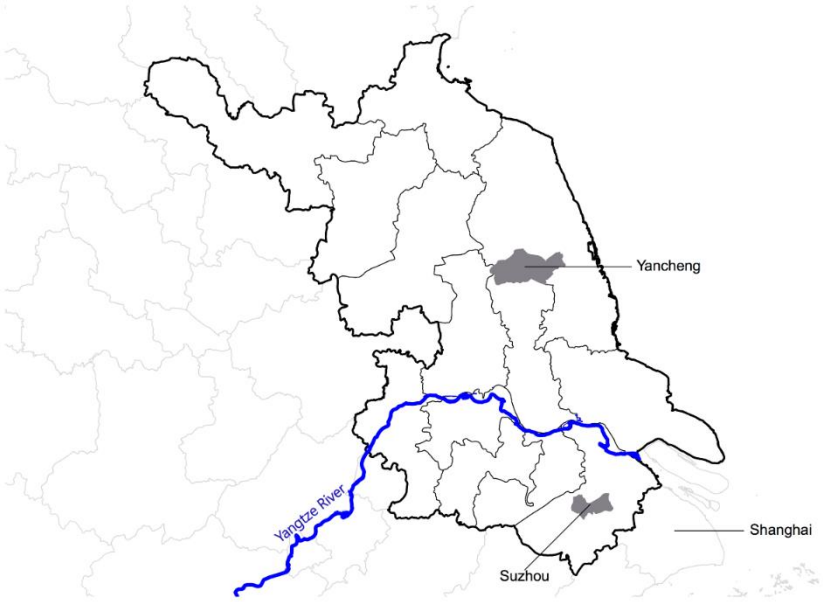
One primary goal in site selection was identifying cities with twice-annual influenza epidemics, given a recent study that identified three regions in China characterized by distinct seasonality: northern provinces (latitude ≥33ºN) experience winter epidemics, southernmost provinces (latitude <27ºN) experience peak activity in the spring and summer, while provinces at intermediate latitudes experience twice-annual epidemic cycles.[1] Appendix Figure 1 indicates the provinces of China in which influenza epidemics occur more than once per year or with irregular seasonality.



Appendix Figure 1: Location of three epidemiological regions in China that are characterized by distinct influenza virus seasonality, as reported by Yu et al.[1] Data were not available for Taiwan and Tibet.

1.2. Jiangsu Province and City Selection.

Jiangsu Province is located in eastern China, just north of Shanghai. Jiangsu is one of the most densely populated Provinces in China, with a population of 79.7 million in 2015. Jiangsu is also one of the more economically developed provinces, ranking second in annual GDP among all provinces, municipalities and autonomous regions in 2015. Jiangsu's rapid urbanization is representative of China's likely future development. Within Jiangsu Province, we selected two cities for study participant enrollment: Suzhou and Yancheng. One primary reason for selecting these two cities within Jiangsu was that the local CDCs had demonstrated capacity to implement rigorous research projects, including previous project collaborations with Suzhou CDC,[2, 3] and another large prospective research study in the community.[4] Suzhou and Yancheng Cities are located on opposite sides of the Yangtze River (Appendix Figure 2), and have distinct geographic and demographic characteristics (Appendix Table 1).



Appendix Figure 2: Location of Yancheng and Suzhou Cities in Jiangsu province, and the Yangtze River

Suzhou (latitude 31°N) is near Shanghai and is considered to be the most economically advanced city in Jiangsu Province. In 2015, the National Bureau of Statistics reported that

Suzhou had the seventh highest annual GDP among all cities in China. Suzhou's urban population grew at a rate of 6.5% between 2000 and 2014, representing the most rapid growth among cities with populations greater than 5 million. At the end of 2015 the resident population of Suzhou was 10.6 million, including 4 million migrants.

Located in northeastern Jiangsu Province, Yancheng (latitude 33°N) has the largest geographic area in Jiangsu. Historically, Yancheng was an agricultural hub. However, the establishment of a Kia Motors plant in 2002 increased Yancheng's industrial production and employment. Nonetheless, according to the Yancheng Bureau of Statistics, young people have continued to leave Yancheng in the past decade to seek work elsewhere, increasing the relative number of older adults in the area.

Appendix Table 1: Land size, population and GDP of Jiangsu Province, Suzhou City and Yancheng City

2015 Snapshot	P. R. China[1]	Jiangsu Province[2]	Suzhou City[3]	Yancheng City[4]
Population, millions	1,374.6	79.7	10.6	7.2
Urban population[5]	56.1%	66.5%	74.9%	60.1%
GDP per capita, RMB	49,351	87,995	136,300	58,299
GDP growth % (yoy)	6.9%	8.5%	7.5%	10.5%
Annual disposable income per capita, RMB				
Overall	21,966	29,500	43,000	22,400
Urban	31,195	37,200	50,400	28,200
Rural	11,422	16,300	25,700	15,700

yoy: year on year. 1 USD = 6.3 RMB in 2015.
1 Source: Statistical Communiqué of National Economic and Social Development (2015)
2 Source: Statistical Communiqué of Jiangsu Province on Economic and Social Development (2015)
3 Source: Statistical Communiqué of Suzhou City on Economic and Social Development (2015)
4 Source: Statistical Communiqué of Yancheng City on Economic and Social Development (2015)
5 Figures on urban population were extracted from the same source as other figures except for that of Suzhou, which is from a report "Analysis on Suzhou's Population and Employment Status in 2015" issued by Suzhou City Bureau of Statistics. Urban residents are defined as those who live and work in cities/towns, while rural residents are those who live in villages and farms for their livelihood.

1.3. Selection of districts and age groups within cities.

We selected two districts, one urban/semi-urban and one semi-urban/rural, within each city to represent older adults living in these two different types of communities. In China, urban areas are traditionally defined as those where the majority of residents live in cities and towns; in rural areas, the majority of residents live in villages and farming is the primary source of income and livelihood. The areas on the outskirts of cities and towns are referred to as semi-urban areas. As both Suzhou and Yancheng are economically developing areas, it is difficult to classify any district as purely urban or purely rural. Indeed, traditional rural areas have been converted into manufacturing zones, such that the GDP per capita is now similar between historically rural and urban areas within these cities. Nonetheless, the living conditions and lifestyle of older adults living in historically rural areas remain different from those in urban communities, and their annual disposable income is lower (Appendix Table 2).

We examined the age distribution of residents in each city using population registries, identifying a total of more than 1 million older adults in each city, with the majority in the age range 60-69 years (Appendix Table 3). We aimed to enroll equal numbers of older adults in three age strata: 60-69 years, 70-79 years, and 80-89 years old. By oversampling from the older age groups, which are frequently underrepresented in research studies, we aimed to examine the incidence of respiratory infections and illnesses with increasing age. Over-enrolling adults aged >70 years will also improve our ability to examine issues of frailty and functional decline. We decided not to enroll adults ≥90 years of age because of the very small number in this group locally (<1 or 2% of adults aged ≥60 years). However, we expect to gain information on those with very advanced age as the cohort ages during the study period.

Appendix Table 2: Land size, population and GDP of four district study sites in Suzhou City and Yancheng City

2015 Snapshot	P. R. China[1]	Suzhou City		Yancheng City	
		Xiangcheng[1]	New District[2]	Tinghu[3,6]	EDZ[4]
Land Area, square km	9,598,077	439	258	732	200
Population, millions	1,374.6	0.73	0.59	0.70	0.20
Classification[5]	-	Semi-urban	Urban	Rural	Semi-urban
GDP per capita, RMB	49,351	83,000	174,000	63,685	N.A.
GDP growth % (yoy)	6.9%	7.4%	8.0%	10.3%.	11.5%
Annual disposable income per capita, RMB[5]					
Urban	31,195	46,100	48,200	33,400	29,900
Rural	11,422	25,400	25,500	17,900	16,200

EDZ: Economic Development Zone. yoy: year on year. N.A.: Data not available. 1 USD = 6.3 RMB in 2015.
[1] Source: Statistical Communiqué of Suzhou Xiangcheng District on Economic and Social Development (2015)
[2] Source: Statistical Communiqué of Suzhou New District on Economic and Social Development (2015)
[3] Source: Statistal Communiqué of Yancheng Tinghu District on Economic and Social Development (2015)
[4] Source: Speech of Ma Junjian at a work meeting of the CCP Working Committee of Yancheng EDZ District, delivered on 26 Jan., 2015, available at http://www.ycedz.com/zt1/news1.asp
[5] Urban residents are defined as those who live and work in cities/towns, while rural residents are those who live in villages and farm for their livelihood. Semi-urban refers to areas in a transitional stage towards urbanization.
[6] Tinghu and EDZ are both districts of Yancheng city. However, the sub-site in Tinghu was recently annexed into the Tinghu district (urban) but still has rural characteristics.

Appendix Table 3: Age distribution of older adults ≥60 in the cities of Suzhou and Yancheng, Jiangsu province, China.

Age group (years)	**Suzhou[1]**		**Yancheng[2]**	
	n	**%**	**n**	**%**
60-69	897,000	56%	704,000	55%
70-79	456,000	29%	424,000	33%
80-89	212,000	13%	131,000	10%
≥90	27,000	2%	14,000	1%
Total	1,592,000		1,273,000	

[1] Source: Suzhou Bureau of Civil Affairs, 2014.
[2] Source: 6th National Census, 2010.

2. Screening and enrolment procedures

2.1. Common recruitment methods across Districts

We aimed to screen and enroll adults in multiple settings to establish a cohort that was representative of the underlying populations of community-dwelling older adults within each district and for all age strata. We wished to avoid enrolling a convenience sample that would exclude homebound or chronically ill older adults. All district study sites used a combination of the following three recruitment methods and settings. First, study sites generated a list of registered older adult residents in the district and then approached potential participants either by door-to-door visit or by telephone contact in order to conduct screening and extend study invitations. Second, we aimed to enroll up to 20% of participants through recruitment in medical settings where potential participants obtain medical or preventive care. Third, for up to 10% of participants enrolled in each location, we allowed enrolment of older adults via personal referrals by other participants. The exact combination of approaches varied slightly by location based on local factors and input from local CDC staff involved in the study.

2.2. Recruitment in two Suzhou Districts

In the urban New District of Suzhou, we aimed to enroll half of the participants from the community and the other half from medical settings and community centers. First, we generated a recruitment list of 200 older adults randomly selected from all residents in each age group in two sub-districts (Xinsheng and Wanfeng). Field research staff attempted to screen everyone on the list by telephone or home visit. If the required number of older adults in an age group was not reached after the first round, an additional 50 older adults from that age group were randomly selected for screening, until a minimum of 50% from each age group was enrolled. Following preliminary screening and initial consent conducted primarily by telephone, potential participants were invited to the community health centre or its subsidiary stations to complete the remaining screening and enrollment procedures. Alternatively, we recruited participants in medical settings, by approaching adults attending outpatient clinic appointments or local chronic disease care programs for screening.

In the semi-urban Xiangcheng district of Suzhou, we also aimed to enroll about half of the participants through direct community invitation. First, after analyzing the age composition

of its 62 village-level administrative groups, field research staff included 11 village groups that were in proximity of the village health clinic where the recruitment was held in the initial recruitment list. The list consisted of 232 potential participants aged 60-89 years, with 135 aged 60-69 years, 68 aged 70-79 years, and 29 aged 80-89 years. Our field staff then worked with local village officials, who visited all potential participants in the selected village groups to invite them to a health service centre for screening. After the initial recruitment list was exhausted, the on-site manager selected additional village groups which were in proximity of the village health clinic and had a higher number of older adults aged 60-89 for further screening. Recruitment in medical settings was conducted simultaneously, and potential participants who sought acute medical services or chronic disease follow-up visits were also approached and screened.

In both the New District and Xiangcheng, due to an insufficient number of older adults aged 80-89 years, we reduced the enrollment goal for participants aged 80-89 years from 125 to 100.

2.3. Recruitment in two Yancheng Districts

In the rural Tinghu district of Yancheng, we selected 10 out of 13 communities in the Yongfeng county, with the three communities excluded due to absence of village clinics, ongoing relocation or remoteness. In each of the communities, a target of 50 participants was set with equal weight of the three age groups. Recruiters started from one village group close to the community committee office and moved to the next village group until the target was met. In the semi-urban Economic Development Zone (EDZ) district, we selected 4 out of 9 communities with a target total of 250 participants. Similarly, recruiters started from one village group and moved to the next village group until the target set for this community was met. Study staff were also allowed to approach visitors of outpatient clinic within the recruitment village groups. Recruitment outside the originally selected communities was permitted after all older adults aged 60-89 years were approached. In Yancheng, field investigators recruited the majority of participants by door-to-door visits.

2.4. Assessment of eligibility to enrolment

Older adults were approached and screened with a standardized form (Annex 1 – Screening Interview) to assess their eligibility to enrolment. Older adults were first asked for their willingness to be screened ('Approached'), their eligibility were then assessed with the list of inclusion and exclusion criteria as described in the manuscript ('Screened'), and those individuals who were deem eligible ('Eligible') were then invited to provide informed consent if they agreed to participate in the study ('Enrolled'). In particular, investigators assessed the cognitive function of the older adults using the standardized Mini-Cog^{TM} tool. Short-term memory was assessed with a three-item word recall. Older adults who could recall all three items ('ball', 'car', and 'man') immediately after they were told and again after a two minute delay (during which other screening questions were asked) were considered eligible. Older adults were excluded if they could not recall all three items initially (after up to 3 attempts), or recalled all items initially but none of them after the delay. If older adults recalled 1 or 2 items after delay, they were asked to complete a standard clock-drawing test (CDT). Older adults who successfully drew the clock remained eligible, while those who could not were excluded. To accommodate the small number of older adults who were unable to hold a pencil due to disability, they were permitted to give step-by-step verbal instructions to staff on how to complete the clock-drawing test.

In addition to screening for eligibility, we had also collected basic information from screened older adults such as age, sex, self-rated health status, knowledge on influenza vaccine and influenza vaccination history in past five years to assess potential differences between screened and enrolled individuals.

Appendix Table 4 below lists the final numbers of older adults approached, screened, eligible, and enrolled in each of the four districts in Suzhou and Yangcheng. Across all study districts, 84% of the older adults approached agreed to complete the screening interview. The response rate was highest (97%) in Xiangcheng and lowest (61%) in New District, both in Suzhou.

Appendix Table 4: Summary of numbers of older adults approached, screened, eligible, and enrolled into the study

	Suzhou			Yancheng			Total
	Xiangcheng (Semi-urban)	**New District** (Urban)	***Sub-total***	**Tinghu** (Rural)	**EDZ** (Semi-urban)	***Sub-total***	
A. Approached	455	648	1103	825	352	1177	2280
B. Screened	440	393	833	767	314	1081	1914
C. Eligible	400	376	776	534	263	797	1573
D. Enrolled	395	376	771	502	259	761	1532
Response (%)	97%	61%	76%	93%	89%	92%	84%
Eligible (%)	91%	96%	93%	70%	84%	74%	82%
Enrolled (%)	99%	100%	99%	94%	98%	95%	97%

Note: Response (%) = B/A, Eligible (%) = C/B, Enrolled (%) = D/C

Appendix Table 5 and Appendix Table 6 compare the characteristics of the enrolled participants with individuals who were approached but not enrolled in Suzhou (1,103 approached, 771 (70%) enrolled, 332 (30%) not) and Yancheng (1,177 approached, 761 (65%) enrolled, 416 (35%) not) respectively. Older adults who were enrolled had significantly higher self-rated health status and were more likely to have heard of the influenza vaccine than older adults who were approached but not enrolled.

Appendix Table 5: Characteristics of enrolled older adults vs older adults approached but not enrolled in Suzhou.

	Enrolled (N=771)		**Approached but not enrolled[1]** (N=332)		**P value**
	n	(%)	n	(%)	
Sex					
Male	335	(43.5%)	142	(42.8%)	0.89
Female	436	(56.5%)	190	(57.2%)	
Age group, in years					
<60	0	0	5	(8.1%)	<0.01
60-69	257	(33.3%)	14	(22.6%)	
70-79	270	(35.0%)	19	(30.6%)	
80-89	244	(31.6%)	23	(37.1%)	
≥90	0	(0)	1	(1.6%)	
Self-reported health status					
Excellent	16	(2.1%)	0	0	<0.01
Very good	194	(25.2%)	9	(16.7%)	
Good	361	(46.8%)	13	(24.1%)	
Fair	190	(24.6%)	25	(46.3%)	
Poor	10	(1.3%)	7	(13.0%)	
Heard of influenza vaccine					
Yes	181	(23.5%)	6	(11.1%)	0.05
No	590	(76.5%)	48	(88.9%)	
Ever received influenza vaccine					
Yes	10	(1.3%)	1	(1.9%)	0.53
No[2]	761	(98.7%)	53	(98.1%)	

[1] Older adults approached but not enrolled were excluded at multiple stages. Therefore, the total numbers in each category do not always add up to 332, but reflect the numbers of the approached/screened who have responded to the specific question.
[2] Older adults who have not heard of influenza vaccine were assumed to have never received influenza vaccine.

Appendix Table 6: Characteristics of enrolled older adults vs. older adults approached but not enrolled in Yancheng

	Enrolled (N=761)		**Approached but not enrolled[1]** (N=416)		**P value**
	n	(%)	n	(%)	
Sex					
Male	345	(45.3%)	185	(44.5%)	0.82
Female	416	(54.7%)	231	(55.5%)	
Age group, in years					
<60	0	0	26	(8.1%)	<0.01
60-69	230	(30.2%)	84	(26.3%)	
70-79	264	(34.7%)	100	(31.3%)	
80-89	267	(35.1%)	105	(32.8%)	
≥90	0	(0)	5	(1.6%)	
Self-reported health status					
Excellent	11	(1.4%)	1	(0.4%)	<0.01
Very good	122	(16.0%)	5	(1.9%)	
Good	220	(28.9%)	28	(10.7%)	
Fair	364	(47.8%)	150	(57.5%)	
Poor	44	(5.8%)	77	(29.5%)	
Heard of influenza vaccine					
Yes	163	(21.4%)	10	(3.8%)	<0.01
No	598	(78.6%)	251	(96.2%)	
Ever received influenza vaccine					
Yes	3	(0.4%)	1	(0.4%)	0.98
No[2]	758	(99.6%)	260	(99.6%)	

[1] Older adults approached but not enrolled were excluded at multiple stages. Therefore, the total numbers in each category do not always add up to 416, but reflect the numbers of the approached/screened who have responded to the specific question.
[2] Older adults who have not heard of influenza vaccine were assumed to have never received influenza vaccine.

Appendix Table 7 to Appendix Table 9 summarise the reasons for non-enrollment of 748 older adults who were approached but not enrolled. These 748 individuals were classified into three categories by their stages of exclusion: 366 older adults who refused screening (Appendix Table 7), 341 older adults who were deemed ineligible (Appendix Table 8), and the small percentage (41/1573, 2.6%) of older adults who refused to give informed consent after they were determined to be eligible for the study (Appendix Table 9).

Appendix Table 7: Reasons approached older adults refused screening

Approached but refused screening	**Suzhou** (N=270)		**Yancheng** (N=96)		**Total** (N=366)	
	n	(%)	n	(%)	n	(%)
Too busy	10	(3.7%)	4	(4.2%)	14	(3.8%)
Time is inconvenient	7	(2.6%)	1	(1.0%)	8	(2.2%)
Not feeling well	3	(1.1%)	1	(1.0%)	4	(1.1%)
Hearing impairment	4	(1.5%)	22	(22.9%)	26	(7.1%)
Other communication impairment	3	(1.1%)	6	(6.3%)	9	(2.5%)
Family member objects	28	(10.4%)	2	(2.1%)	30	(8.2%)
Local recruitment facility staff objects	2	(0.7%)	0	(0)	2	(0.5%)
Not interested	148	(54.8%)	17	(17.7%)	165	(45.1%)
Other	65	(24.1%)	43	(44.8%)	108	(29.5%)

Appendix Table 8: Reasons screened older adults were excluded as ineligible

Screened but ineligible[1]	**Suzhou** (N=57)		**Yancheng** (N=284)		**Total** (N=341)	
	n	(%)	n	(%)	n	(%)
Age restriction (not in the range 60-89y)	6	(10.5%)	31	(10.9%)	37	(10.9%)
Residence restriction						
Does not live in study city	2	(3.5%)	7	(2.5%)	9	(2.6%)
Not intending to live in study city for the next 2 years	0	(0)	21	(7.4%)	21	(6.2%)
Does not have a landline or cellular/mobile	9	(15.8%)	36	(12.7%)	45	(13.2%)
Medical exclusion						
Bleeding disorder	6	(10.5%)	1	(0.4%)	7	(2.1%)
Anticoagulant use	3	(5.3%)	11	(3.9%)	14	(4.1%)
Cognitively impaired						
Recalled 0 of 3 words immediately	10	(17.5%)	56	(19.7%)	66	(19.4%)
Retained 0 of 3 words	14	(24.6%)	43	(15.1%)	57	(16.7%)
Retained 1 or 2 of 3 words but failed clock drawing task	7	(12.3%)	78	(27.5%)	85	(24.9%)

[1] Individuals were screened for each criterion and excluded immediately if fail to meet the criterion. No individuals were excluded due to having a history of severe reaction to influenza vaccination that required medical attention.

Appendix Table 9: Reasons eligible older adults refused enrolling into the study

Eligible but refused enrollment	Suzhou (N=5)		Yancheng (N=36)		Total (N=41)	
	n	(%)	n	(%)	n	(%)
Too busy	0	(0)	1	(2.8%)	1	(2.4%)
Time is inconvenient	0	(0)	0	(0)	0	(0)
Not feeling well	1	(20.0%)	1	(2.8%)	2	(4.9%)
Hearing impairment	0	(0)	2	(5.6%)	2	(4.9%)
Other communication impairment	0	(0)	1	(2.8%)	1	(2.4%)
Family member objects	0	(0)	2	(5.6%)	2	(4.9%)
Not interested	1	(20.0%)	24	(66.7%)	25	(61.0%)
Other	3	(60.0%)	5	(13.9%)	8	(19.5%)

After written informed consent was obtained, study staff administered part 1 of a two-part structured questionnaire (described in the following section) to collect baseline information on basic demographics, health status and contact details for surveillance activities, while trained phlebotomists collected 5-10ml of blood in tubes with clot activators for the assessment of influenza infection history (described in section 8.1). A longer part 2 of the two-part structured questionnaire, which collected more in-depth baseline information, was completed immediately or at a follow-up meeting with the participant within 1-3 months after enrolment.

3. Baseline survey instrument

A baseline survey instrument was designed to collect sufficient baseline information to measure potential changes in frailty and functional status over time, as well as other relevant demographic and clinical information. The cohort study also aimed to establish the feasibility of potential vaccination trials in this location, and was an opportunity to pilot questions which may be useful in future studies.

The baseline survey instrument was divided into two parts to streamline the enrolment process. The first part (Annex 2 – Enrolment Interview 1) was performed immediately after written informed consent was obtained from each participant and included basic

demographic, household, and health information. A detailed health history was not included in this section, but participants were asked a single-item summary question to identify any underlying medical conditions. Basic questions about living conditions and education level, as well as several questions from standardized measures of frailty and functional status were also asked. Height, lower leg length and weight were measured. Lower leg length was measured as it does not shrink with age; hence, it is a better proxy for pre-shrinkage height in older adults.[5]

A longer structured questionnaire administered by study staff (Annex 3 – Enrolment Interview 2) collected more in-depth information on participants' functional status, cognitive function, life history, socio-economic status, health conditions, and attitudes toward influenza vaccination. This was completed immediately or at a follow-up meeting with the participant within 1-3 months after enrolment. Participants were randomly assigned to receive either version A or version B of the Enrolment Interview 2 questionnaire, with approximately half of the participants completing each version. Eventually, 1,506/1,532 (98%) participants completed Enrollment Interview 2, with 758/1506 (50%) administered version A, and the remaining (748/1506, 50%) particpants version B of the Enrollment Interview 2 questionnaire.

Both version A and version B of Enrolment Interview 2 included a common set of items, and additional extended questions specific to that version (Appendix Table 10). All participants were surveyed on functional status (more details below), cognitive function proxy by SMMSE (more details below), life history as older adult after 60 years of age (more details below), socioeconomic status (personal annual income), general health (5-graded self-rated health and smoking history), chronic medical conditions (diagnosed conditions, current medications including steroids, unintentional weight loss, falls, hospitalisations and self-rated change in overall health), and knowledge, attitudes and practices (KAP) on influenza virus and vaccination. In addition, 758 participants who were administered the version A of Enrolment Interview 2 answered extended questions on socio-economic status (measures of material well-being and self-rated perceptions of social status), general health (self-rated health on a scale of 0-100, exposure to second-hand smoking and pneumococcal vaccination), and an additional section on life history before 60 years of age (more details

below). On the other hand, 748 participants who were administered the version B of Enrolment Interview 2 answered extended questions and with more extensive scaling on functional status and depression (more details below).

All participants were surveyed regarding functional status that included questions from the Groningen Activity Restriction Scale (GARS)[8] and the Groningen Frailty Index (GFI).[9] To evaluate the utility of additional items to measure functional status and any changes in the level of an individual participant's frailty over the course of the study, Version B included additional questions on self-rated physical and mental health in the past 30 days from the BRFSS 2014 Questionnaire[7], and more extensive questions on activities of daily living and other measures which could be used to judge frailty and to compare this cohort's participants with studies done in other settings using standardized frailty measures. This extended version incorporates scaling of difficulty for activities that can be done independently but still may be a challenge for the older person, using the response format featured in the GARS, and included all the items in this scale plus overlapping items with other prioritized indices. Version B also included additional standardized items to evaluate the frequency of depressive symptoms in the past two weeks, which were adapted from the Center for Epidemiologic Studies Depression Scale (CES-D).[10]

Although all participants had had basic cognitive screening prior to enrolment to ensure that they did not show signs of dementia or significant cognitive impairment, which would exclude them from participating in the study, in Enrollment Interview 2 all participants also completed the Standardized Mini-Mental State Examination (SMMSE) for more comprehensive evaluation of their cognitive function.[6]

As part of the section on life history as older adult after 60 years of age, all participants were asked about their current employment status, level of activity and social engagement using questions adapted from the NIA Health Retirement Study and the US BRFSS 2014 Questionnaire.[7] The version A also included an additional section on life history before 60 years of age, which was sub-divided into three periods of life: as child, as young adult between 18-34 years old and as middle-aged adult between 35-59 years old. In this additional

section, information on city of birth or residence (for classification of urban versus rural environment), mobility, and parents' and own's occupations were collected.

Appendix Table 10: Structure of the main baseline assessment, done as soon as possible after enrolment. Participants were randomly allocated to two alternative instruments with partial overlap.

Section[1]	Enrollment 2A Interview	Enrollment 2B Interview
A	Administrative Info.	Administrative Info.
I	Functional Status (Brief)	-
J	-	Functional Status (Extended, incl. Brief)[2]
L	SMMSE	SMMSE
B	Life History Child to Adult	-
C	Life History Older Adult	Life History Older Adult
D	SES (Brief)	SES (Brief)
E	SES (Extended)	-
F	-	General Health (Brief)
G	General Health (Extended, incl. Brief)[2]	-
H	Chronic Disease	Chronic Disease
K	Influenza Vaccination KAP	Influenza Vaccination KAP

SES: Socioeconomic status. SMMSE: Standardised Mini-Mental State Examination. KAP: Knowledge, Attitude and Practice.

[1] The sections were not in alphabetical order since sections were rearranged at the start of the study to minimize response fatigue in subjective scales.

[2] The Extended version of Functional Status in Enrollment 2B Interview included all questions in the Brief version in Enrollment 2A Interview; similarly, the Extended version of General Health in Enrollment 2A Interview included all questions in the Brief version in Enrollment 2B Interview.

4. Re-assessment instruments during follow-up and at withdrawal

Halfway through every study year around spring, a short questionnaire on health status and change (Annex 4 – Half Year Follow-up Interview) is administered in all participants, in-person for the sub-group of participants who are also providing blood samples mid-year, and over telephone in the remaining participants. It includes self-rated general, physical and

mental health in the past 30 days from the BRFSS Questionnaire, and change in overall health in the past 6 months.

An annual re-assessment is completed through a staff-administered questionnaire (Annex 5 – Annual Reassessment) at the end of each study year (which is the start of the following year). In addition to assessing whether any new health conditions have developed or existing conditions have worsened, particularly in the setting of influenza virus or RSV infection, the annual reassessment is designed to assess changes in socioeconomic status, functional status, the level of social engagement, and self-reported health status. Preliminary data from the baseline survey instrument was examined to determine which questions were most useful in discriminating between participants in the study, and therefore annual reassessment questionnaires only included a subset of questions from the baseline survey instrument. Participants are asked about any recent hospitalizations or other changes to their medical condition, including falls, unintentional weight loss, and other information which may not have been captured during active respiratory illness surveillance. Questions adapted from the GARS, GFI, US BRFSS Questionnaire and CES-D are repeated in the annual reassessment, which can be compared with the baseline assessment to examine whether a decline in functional status and worsening frailty can be measured. Several questions regarding knowledge, attitude and practice towards influenza vaccines are also asked, as a follow-up to the KAP study conducted as part of the initial assessment. A full SMMSE was not administered at the first annual reassessment in 2016 but will be administered at the second annual reassessment to all participants in 2017; these questions can be compared to baseline SMMSE scores.

When a participant seeks to withdraw from the study at any point of the study, a short questionnaire on time and reasons of withdrawal including hospitalizations (Annex 13 – Participant Withdrawal Form) is administered before he/she exits from the cohort to record reasons for withdrawal, allowing us to assess whether the withdrawal might be related to worsening of health or disability.

5. Active identification of acute respiratory illnesses and hospitalizations

Enrolled participants are followed up with active surveillance activities weekly throughout the year to identify acute respiratory illnesses, which permits home visits for collection of nasal and throat swabs during an acute illness for testing by rRT-PCR to confirm influenza virus and RSV infections, and completion of illness surveys on symptoms severity, cognitive function, subjective health status, medical care and daily activities during the early and late stage of the acute illness.

Participants are asked to contact study staff directly whenever they feel sick with any of these symptoms. In addition, study staff telephone participants weekly to monitor the occurrence of illness symptoms in the past week ('In the past 7 days have you felt ill or sick?'), and remind participants to call study staff directly if they become ill. During periods of heightened local influenza circulation, study staff may call participants twice weekly. Study staff may also conduct home visits for older adults who are unable to use the phone (and lack caregivers who can answer for them) or have hearing loss or another disability that preclude telephone surveillance. Active surveillance efforts are conducted throughout the year, except for Chinese New Year because for cultural and logistical reasons we are unable to conduct active surveillance and home visits during this festival, which takes place over approximately 1 week at the end of January or early to mid February.

When an illness is reported by a participant, trained study staff completes brief screening questions to identify symptoms and the illness onset date (Annex 8 – Symptom Screening Log). Acute illness is defined by two or more of the following symptoms: fever (feverishness, chills, or elevated temperature ≥37.8°C), runny nose, worsened shortness of breath, sore throat, cough, body or muscle aches and pain, and headache. Depending on the number of symptoms identified, the illness onset date and whether the illness has already been resolved, different illness surveys with or without respiratory specimen collection are initiated (next section).

In Suzhou, active surveillance started from 28 December 2015 before the completion of participant recruitment. In Yancheng, active surveillance started from 11 January 2016, approximately one month after the completion of participant recruitment. From 22 February

to 31 March 2016 because of evidence of increased influenza activity in the community, Suzhou increased the contact frequency to twice a week while Yancheng maintained as once a week.

Counts of newly developed respiratory illnesses are summarized on a weekly basis from Sunday to Saturday, coincides with the US CDC Morbidity and Mortality Weekly Report (MMWR) surveillance weeks. During each surveillance week, at least two contacts per day for three consecutive days are attempted for each participant until the study staff reaches the participant ('success'). Additional effort including home visits may be conducted to reach the participants if contact through surveillance calls fails. If a participant is not reached despite multiple efforts, the contact is considered as 'failure' for that week. Participants currently experiencing an acute illness are 'ineligible' for active surveillance, and is resumed the week after the participant's illness resolves.

Appendix Table 11 shows the contact rates in weekly active surveillance by age group for the first year of the study (until 3 Septebmer 2016). For both Suzhou and Yancheng, on average we were able to reach the participants and confirmed their health status in over 90% of the person-weeks of follow-up.

Starting from April 2017, we are also actively identifying hospitalizations in the past month on a monthly basis regardless of whether it is related to an acute illness (Annex 14 – Monthly Hospitalization Surveillance Form). We collect information on the admission date, type of medical care, duration and reasons of hospitalization.

Appendix Table 11: Contact rates in weeks of active surveillance by age group for the first year of the study.

Study site and age group, in years	**Person-week of follow-up**, n (row %)			
	Success	**Ineligible**	**Failure**	**Total**
Suzhou				
60-69	9607 (94%)	108 (1%)	484 (5%)	10199
70-79	8510 (93%)	55 (1%)	578 (6%)	9143
80-89	6544 (91%)	53 (1%)	588 (8%)	7185
Total	24661 (93%)	216 (1%)	1650 (6%)	26527
Yancheng				
60-69	8141 (96%)	50 (1%)	259 (3%)	8450
70-79	8118 (96%)	59 (1%)	275 (3%)	8452
80-89	8120 (96%)	37 (1%)	259 (3%)	8416
Total	24379 (96%)	146 (1%)	793 (3%)	25318

Contact attempts for each MMWR surveillance week are classified into 3 categories: Participants who were successfully reached by at least one call during the week ('Success'); participants who were currently experiencing an acute illness and being monitored and thus were not eligible for active surveillance for identification of new illness ('Ineligible'); and participants who could not be reached during the week (if any) ('Failure').

6. Monitoring of acute illnesses

Once an acute illness is identified, depending on the number of symptoms identified, the illness onset date and whether the illness has already been resolved, different surveillance activities (illness surveys with or without respiratory specimen collection) will be completed (Appendix Table 12).

A home visit is scheduled for participants with at least two illness symptoms and illness onset within the prior 7 days for the collection of respiratory specimens and the administration of the Acute Illness Interview (Annex 9). At the home visit, trained study staff collect a respiratory specimen using mid-turbinate nasal and oropharyngeal swabs and ask the participant to describe the possibility of household transmission, symptom severity at present, medical care including prescription medication use and any disruption to normal activities

due to the illness (Annex 9 – Acute Illness Interview). After screening for eligibility (e.g. literacy, vision impairment, writing difficulty) for completing a symptom diary (Annex 11 – Symptom Diary), participants are asked to describe their illness symptoms for each subsequent day during their illness using the symptom diary for each illness episode. Approximately 10 days after illness onset, healthcare workers telephone participants to repeat most questions in the Acute Illness Interview except symptom severity when the participant is most ill, and also ask if the illness has resolved (Annex 10 - Illness Follow-up Interview). Participants who have not recovered at this time are telephoned every 3 days up to four times or until an illness resolution date is identified.

If a participant is reported dead during the acute illness, we attempt to identify the date and cause of death (Annex 12 – Death Record, refer to section 7.3 below). In a subset of participants, information regarding hospitalization (if any) for acute illness (Annex 7 – Hospital Case Report) is also abstracted (refer to section 7.2 below).

Appendix Table 12: Study activities once an acute illness is identified.

	Respiratory Specimen Collection	**Acute Illness Interview + Symptom Diary**	**Illness Follow-up Interview**
Onset in prior 7 days			
Still sick	x	x	x
Illness resolved	x		x
Onset in prior 8-10 days			
Still sick		x	x
Illness resolved			x
Onset in prior 11 days at least			
Still sick			x
Illness resolved			x

7. Review of medical charts, hospitalization records, and death records

7.1. Review of medical charts

The medical chart review is designed to obtain comprehensive and accurate medical histories for enrollees to meet two primary objectives: 1) to assess the association of chronic medical conditions and risk of influenza virus and RSV infections among older adults in the cohort; and 2) to provide validating data for baseline medical status that was assessed during the enrolment interview. Supplementing the self-reported information with data collected from medical records is considered valuable because most study participants reported having received very little if any formal education, which may have limited their understanding of chronic medical conditions. A team of Western and Chinese epidemiologists and clinicians developed a data abstraction tool (Annex 6 – Medical Chart Review) with sections covering demographic information, chronic medical conditions and hospitalizations.

Most adults in China do not have designated primary care doctors, and there is not yet a national system to record individuals' comprehensive health information. However, both Suzhou and Yancheng CDCs have established community health information systems for chronic conditions to record basic primary health data for community-dwelling residents in their cities. This electronic system records the diagnosis and management of four chronic conditions: hypertension, diabetes, stroke and cancers. In consultation with the local CDCs, the study team determined that this system is the most reliable data source available for study participants' medical records in both Suzhou and Yancheng. However, the electronic community health information system is still being constructed from existing paper records, limiting the completeness of data available. When study staff identify discrepancies between self-report health information and data in the medical records, further assessment may be conducted to verify records (e.g. with the paper records) on a case-by-case basis.

7.2. Hospital Chart Review

The hospital chart review is designed to 1) characterize influenza illness requiring hospitalization among those with laboratory-confirmed influenza infection during the study and 2) assess all hospitalizations for acute illness during the study period, focusing on clinical diagnoses and duration of hospitalization. A team of Western and Chinese epidemiologists

and clinicians developed a data abstraction tool (Annex 7 – Hospital Case Report) with sections covering demographic information, medical and vaccination history and information on present illness requiring hospitalization including clinical diagnoses, laboratory testing results, interventions (e.g., oxygen support, mechanical ventilation), treatments, and duration of hospitalization (Annex 7 – Hospital Case Report). We intend to access participants' electronic health records in major hospitals in Suzhou and Yancheng, subject to hospital permission and a separate participant's written consent for accessing their medical records in these hospitals.

7.3. Death Record Review

To identity influenza-related deaths and to obtain accurate data on date, and primary and secondary causes of death from study participants who passed away during the study period, we designed a data abstraction tool (Annex 12 – Death Record) which consists of two sections: (1) data to be collected from a family member interview and (2) data to be abstracted from an official report. The death of a study participant is first reported to study staff by family members during the weekly active surveillance calls, and within 30 days study staff arrange an in-person interview with the family members. They also verify the death with any official records including death certificates and hospital records. Information solicited from family members includes interviewee's relationship to the participant, prior illness and/or hospitalization, place, date and causes of death. Data abstracted from offical records includes the place, date, nature and causes of death, whether an autopsy has been performed, and information regarding the person(s) who pronounced and certified the death.

8. Collection, storage and laboratory testing of biological specimens

The primary outcome measures are influenza virus infections and respiratory syncytial virus (RSV) infections confirmed by rRT-PCR, and serologic evidence of influenza virus infection. All respiratory specimens collected during active surveillance are screened for influenza A/B viruses and RSV by rRT-PCR. For specimens positive for influenza A/B viruses, additional rRT-PCR for influenza A virus subtypes (H1pdm09 and H3), influenza B virus lineages (Victoria and Yamagata), and absolute quantification (expressed in copies/mL) are conducted. A PCR-confirmed influenza virus infection or RSV infection is defined as

positive result tested by RT-PCR (Ct ≤40 for influenza, Ct ≤42 for RSV) on the combined nose and throat swab collected in a home visit during acute illness. Serologic evidence of influenza virus infection is defined as either a ≥4 fold rise in antibody titre between consecutive paired sera, or an increase from HI <10 to HI ≥40 during the same interval, in participants without influenza vaccination prior to the immediate influenza season.

8.1. Collection of blood specimens

Serological specimens are collected at enrolment and every 12 months throughout the study; in addition serologic specimens are collected every six months in a random subset of participants. In each blood draw we collect 5-10 ml blood using vacutainer tubes with clot activators. Phlebotomists use a butterfly needle connected to a vacutainer tube to minimize hemolysis and to reduce the risk of needle stick injury. After collection, the blood tubes are stored in a cool box with at least 2 ice packs immediately, transported and maintained at 2-8℃ en route to the laboratory at the city CDCs.

8.2. Laboratory processing and storage of sera specimens

Upon arrival at the laboratory, aliquot tubes with barcoded labels corresponding to the specimen labels on the blood collection tubes are prepared. After centrifugation, sera derived from clotted blood is aliquoted and stored at -80℃ or, if storage is not available, at -20℃ for less than 1 year.

Each serum specimen will be divided into 3-4 aliquots. Paired serum specimens (before and after influenza seasons) will be tested for antibody responses to vaccine strains and, if available, circulating influenza strains by hemagglutination inhibition (HAI) assays.

8.3. Influenza serology

Sera from the same participants are tested in parallel.[14] Briefly, sera are thawed and treated with receptor-destroying enzyme to removed non-specific inhibitors, then heat-inactiviated at 56℃ for 30 min. The sera are then absorbed with turkey red blood cells to minimize non-specific agglutination. Antibody titers are determined by testing serial two-fold dilutions from 1/10 to 1/1,280 in duplicate, in 96-well microtiter plates with 0.5% turkey erythrocytes

using four hemagglutination units. Any uncertain results are resolved by repeat testing in quadruplicate. Positive and negative control sera are also tested at the same time with virus back-titration performed. Repeated laboratory assays will be done in 10% of the specimens using a separate aliquot for validation.

8.4. Collection of respiratory specimens

At the home visit during illness surveillance, trained study staff collect a respiratory specimen using combined mid-turbinate nasal and oropharyngeal swabs. Respiratory specimens collected during home visits are stored in a cool box with at least 2 ice packs immediately after collection, transported and maintained at 2-8℃ en route to the laboratory at the city CDCs within 24 hours after collection.

8.5. Laboratory processing and storage of respiratory specimens

Upon arrival at the laboratory, aliquot tubes with barcoded labels corresponding to the specimen labels on the swab collection tubes are prepared. After vortexing and removal of the swabs, viral transport medium from each collection tube is aliquoted and stored at -80℃. Each combined nose and oropharyngeal swab specimen is divided into 3-4 aliquots; the first aliquot will be for RT-PCR testing of influenza virus type, subtype/sub-lineage, absolute quantification and RT-PCR testing of RSV, the second aliquot is used for confirmation or further testing (e.g., to identify other respiratory pathogens); additional aliquots are stored for future study uses.

Real-time reverse transcription polymerase chain reaction (rRT-PCR) are completed in a reference laboratory of the local city CDC using United States Centers for Disease Control and Prevention (US CDC) primers, probes, reagents, and protocols for PCR testing of influenza virus type, subtype/sub-lineage and RSV; and in-house reagents and protocols provided by the University of Hong Kong for the absolute quantification of influenza virus. Laboratory assays will be repeated in 10% of the specimens using a separate aliquot for validation.

8.6. Total RNA extraction

The combined nose and oropharyngeal swab specimen is subjected to total RNA extraction using existing extraction systems in the local (Suzhou/ Yancheng) CDCs.

In Suzhou, RNA extraction is performed by the QIAsymphony SP platform (Qiagen, Hilden, Germany) using QIAsymphony Virus/Bacteria Mini kit (Cat # 931036) according to the manufacturer's instructions (protocol Complex200_V6_DSP). A respiratory specimen is first equilibrated to room temperature; 200 µl of respiratory specimen is transferred to a 2ml tube and placed in the tube carrier, where the QIAsymphony SP instrument conducts automatic extraction. 60 µl of RNase-free elution buffer is used for the recovery of nucleic acid.

In Yancheng, RNA extraction is performed by the Ambion MagMax Express (24-well low throughput) platform (Life Technologies, Carlsbad, CA) using Ambion MagMax-96 Viral RNA Isolation Kit (Cat # AM1836) according to the manufacturer's instructions (protocol AM1836v2). Briefly, 150 µl of wash solution 1, 150 µl of wash solution 2, 90 µl of elution buffer, 20 µl of bead mix, 50 µl of the respiratory specimen, and 130 µl of lysis/binding solution is added successively to the processing plate, and the MagMax Express Magnetic Particle Processor conducts automatic extraction. A final volume of 50 µl of nucleic acid per sample is recovered.

In both Suzhou and Yancheng, extracted RNA is kept at -80°C until further processing by real-time PCR for influenza virus or RSV detection.

8.7. Influenza PCR

In both Suzhou and Yancheng, total RNA extracted is tested for influenza virus type and subtype/lineage by real-time RT-PCR using US CDC primers, probes, reagents and protocols. Absolute quantification of influenza A/B virus is conducted using in-house reagents and protocols provided by the University of Hong Kong.[12]

Screening for type A and type B influenza viruses is conducted according to the CDC real-time RT-PCR Protocol for Detection and Characterization of Influenza. Briefly, influenza A

and B virus RNA is detected by one-step real-time RT-PCR using the Ambion AgPath-ID One-Step RT-PCR Kit (Cat # 4387391, Applied Biosystems, Waltham, MA). The primers and probes (CDC Influenza Virus Real-Time RT-PCR Influenza A/B Typing Panel, Cat # FluSS-01, International Reagent Resource (IRR), US CDC) are designed to detect the matrix (M) protein gene of influenza A and B virus. The reaction master mix of the assay includes 12.5 µl 2X PCR Master Mix, 1 µl RT Mix, 0.5 µl 40µM forward primer (final concentration 0.8µM), 0.5 µl 40µM reverse primer (final concentration 0.8µM), 0.5µl 10µM TaqMan probe with FAM dye (final concentration 0.2µM), 5 µl Nuclease free water and 5 µl of RNA to a final volume of 25 µl. The reaction is performed in ABI 7500 system (Applied Biosystems). The cycling condition is as follows: an initial reverse transcription step at 50℃ for 30 min and an enzyme pre-activation step at 95℃ for 10 min, then proceeded to 45 cycles of amplification steps (95℃ for 15 sec denaturation, 55℃ for 30 sec anneal/extension). An experiment run is considered valid when the result for no template controls (NTC) and mock extraction control (MOCK) is negative (without crossing the threshold line for the 40 cycles of reaction), and the result for positive template controls (PTC) is positive and within the expected Ct values. PTC includes Pooled Influenza Positive Control (Cat # VA2716, IRR). All clinical samples should exhibit RP reaction curves that cross the threshold line at or before 35 cycles. A clinical sample is defined as positive for influenza when the reaction curve crosses the threshold line before 40 cycles (i.e. $Ct \leq 40$).

Determination of influenza A virus subtypes and influenza B virus lineages are conducted following similar procedures as in screening (above). Primers and probes for influenza A virus subtyping are provided in CDC Influenza Virus Real-time RT-PCR Influenza A (H3/ H1 pdm09), Subtyping Panel (version 2) (Cat # FluRUO-09, IRR) with Pooled Influenza Positive Control (Cat # VA2716, IRR), and for influenza B virus lineage in CDC Influenza B Lineage Genotyping Panel (Cat # FluRUO-05, IRR) with Influenza B Positive Control (Cat # VA2733, IRR).

For absolute quantification, a reference standard is prepared using pCRII-TOPO vector (Invitrogen, San Diego, CA) containing the corresponding target viral sequences. A series of eight log 10 dilutions equivalent to 1×10^0 to 1×10^7 copies per reaction are prepared to generate calibration curves and run in parallel with the test samples. The primers and probes

are designed to detect the matrix (M) protein of influenza A and B virus. The reaction master mix of the assay includes 12.5µl 2X PCR Master Mix, 1 µl RT Mix, 1 µl primers-probe mix, 5.5 µl Nuclease free water and 5µl of RNA to a final volume of 25µl. If the specimen result is outside the upper limit of the expected range, the extract of the sample is repeated with suitable dilution. The detection limit for this assay is 10 copies per reaction.

8.8. RSV PCR

In Suzhou and Yancheng, screening for respiratory syncytial virus (RSV) is conducted according to the US CDC protocol Real-Time rRT-PCR Assays for Non-Influenza Respiratory Viruses with reagents provided by US CDC.[13] Briefly, RSV type A or B RNA is detected by one-step real-time RT-PCR using the Ambion AgPath-ID One-Step RT-PCR Kit (Cat # 4387391, Applied Biosystems, Waltham, MA). The primers and probes (Division of Viral Diseases, NCIRD, US CDC) are designed to detect the matrix (M) protein of RSV type A or B. The reaction master mix of the assay includes 12.5 µl 2X PCR Master Mix, 1 µl RT Mix, 0.5 µl 50X forward primer, 0.5 µl 50X reverse primer, 0.5 µl 50X probe, 5 µl Nuclease free water and 5 µl of RNA to a final volume of 25 µl. The reaction is performed in ABI 7500 system (Applied Biosystems). The cycling condition is as follows: an initial reverse transcription step at 45℃ for 10 min and an enzyme pre-activation step at 95℃ for 10 min, then proceeded to 45 cycles of amplification steps (95℃ for 15 sec denaturation, 55℃ for 60 sec anneal/ extension). An experiment run is considered valid when the result for NTC is negative, and the result for Viral Template Control (VTC) (Division of Viral Diseases, NCIRD, US CDC) is positive and within the expected Ct values. All clinical samples should exhibit positive RP reaction curves. A reaction is defined as positive when an exponential curve is produced with a sharp increase in fluorescence, with Ct ≤37 generally accepted as true positive. Specimens with $38 \leq Ct \leq 42$ are considered as a weak positive, the result of which will be interpreted with caution and repeat testing may be conducted.

8.9. Additional laboratory testing of sera and respiratory specimens

We plan to perform additional assays on serum specimens focused on biomarkers of disease severity, the hemagglutinin (HA) or neuraminidase (NA) influenza antigens, and antigens specific to RSV or other pathogens. For example, the enzyme-linked lectin assay (ELLA)

may be used to assess neuraminidase inhibition antibody responses. Additionally, some specimens may be stored long-term as part of a specimen bank, but will not be linked to any participant identifiers. These specimens may also be utilized to investigate novel viruses or pathogens, especially to address future pandemic pathogens.

Remaining aliquots of all study respiratory specimens will be sent for banking and storage according to the relevant regulatory requirements in China; no specimens will contain personal identifiers. Respiratory specimens may be utilized in further studies to investigate novel viruses or pathogens, especially to address future pandemic pathogens, or biomarkers of disease severity.

9. Data management in REDCap

While multi-site and multi-domain projects can be efficient and effective for collecting data on hard to reach populations or seasonal data, they are also complex in terms of ensuring standardized data collection across study sites. Furthermore, a clearly defined common data management process, and tools such as data dictionaries and standardized data entry platforms are necessary to ensure that final datasets are developed using methodologically sound data collection, entry, and cleaning processes. As shown below in Appendix Table 13, these requirements included system availability and flexibility, user access and functionality, and the ability to customize the tool to meet the specific needs of the project.

Appendix Table 13: Data Management System requirements.

Criteria	Description
Easily accessible	Secure, web-based data entry
Fast and flexible	Able to quickly pre-populate databases for new studies with common data elements
Multi-level user access	Investigators and study staff have customized person-specific token access
Fully customizable	Ability to format data entry screens to match study-specific questionnaires and other forms
Advanced question features	Auto-validation, branching logic, and stop actions
Data import functions	Data may be imported from external data sources (for example, site electronic medical records)
Survey export functions	Export survey results to common data analysis packages: (e.g. Microsoft Excel, SAS, Stata, R, or SPSS)

Given its flexibility and applicability across study sites, REDCap[11] was used as the centralized data capture system. Developed by Vanderbilt University, with collaboration from a consortium of institutional partners, REDCap is a software toolset and workflow methodology for electronic collection and management of research and clinical trial data.[11] REDCap is a secure, web-based application designed to support data capture for research studies, providing: 1) an intuitive interface for validated data entry; 2) audit trails for tracking data manipulation and user activity; 3) automated export procedures for data downloads to Excel, PDF, and common statistical packages (SPSS, SAS, Stata, R); and 4) procedures for importing data from external sources. Also included are a built-in project calendar, a scheduling module, ad hoc reporting tools, and advanced features, such as branching logic, file uploading, and calculated fields. REDCap is designed to comply with regulations under the Health Insurance Portability and Accountability Act.

The data collection forms (study instruments) and data dictionary were finalized through collaboration between all investigators and identical across sites, and housed as a common project database. For the administration of the study instruments, during the in-person

interview with the participant we enter data directly into REDCap in real-time (via a computer or tablet with a wireless connection). Interviews conducted over telephone calls, active surveillance calls, specimens collection and laboratory results are also tracked and recorded using REDCap. Data are downloaded by the investigators on a regular basis and in real-time for quality assurance checks and tracking, includes checking data for adherence to the common protocol, outliers, and missing or incomplete data in real-time to assist in immediate retification, and checks post-collection for further cleaning and preparation of the final dataset.

The prinicipal investigator (PI) and the study manager designated by PI are responsible for safekeeping of the personal data during and after the study. All data are anonymized and stored in the server located in the University of Hong Kong. Original identities will be kept in a separate file accessible only to the trial manager. Original paper documents (consent form) will be destroyed after retention period of 3 years or per local IRB requirement. None of the subjects' personal information will be revealed in any subsequent research output.

10. References

1. Yu H, Alonso WJ, Feng L, et al. Characterization of regional influenza seasonality patterns in China and implications for vaccination strategies: spatio-temporal modeling of surveillance data. *PLoS Med* 2013; **10**: e1001552.
2. Wang Y, Zhang T, Chen L, et al. Seasonal influenza vaccine effectiveness against medically attended influenza illness among children aged 6-59 months, October 2011-September 2012: A matched test-negative case-control study in Suzhou, China. *Vaccine* 2016; **34**: 2460-5.
3. Zhang X, Zhang J, Chen L, et al. Pneumonia and influenza hospitalizations among children under 5 years of age in Suzhou, China, 2005-2011. *Influenza Other Respir Viruses* 2016.
4. Zhang J, Zhang XF, Huang SJ, et al. Long-term efficacy of a hepatitis E vaccine. *N Engl J Med* 2015; **372**: 914-22.
5. Huang W, Lei X, Ridder G, Strauss J, Zhao Y. Health, Height, Height Shrinkage, and SES at Older Ages: Evidence from China. *Am Econ J Appl Econ* 2013; **5**: 86-121.
6. Molloy DW, Alemayehu E, Roberts R. Reliability of a Standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. *Am J Psychiatry* 1991; **148**: 102-5.
7. United States Centers for Disease Control and Prevention. *Behavioral Risk Factor Surveillance System*. [cited October 19 2016]; Available from: http://www.cdc.gov/brfss/
8. Suurmeijer TP, Doeglas DM, Moum T, et al. The Groningen Activity Restriction Scale for measuring disability: its utility in international comparisons. *Am J Public Health* 1994; **84**: 1270-3.
9. Bielderman A, van der Schans CP, van Lieshout MR, et al. Multidimensional structure of the Groningen Frailty Indicator in community-dwelling older people. *BMC Geriatr* 2013; **13**: 86.
10. Radloff LS. The CES-D scale: A self report depression scale for research in the general population. *Appl Psychol Meas* 1977; **1**: 385-401.
11. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009; **42**: 377-81.

12. Ip DK, Lau LL, Chan KH, et al. The Dynamic Relationship Between Clinical Symptomatology and Viral Shedding in Naturally Acquired Seasonal and Pandemic Influenza Virus Infections. *Clin Infect Dis* 2016; **62**: 431-7.
13. Fry AM, Chittaganpitch M, Baggett HC, et al. The burden of hospitalized lower respiratory tract infection due to respiratory syncytial virus in rural Thailand. *PLoS One* 2010; **5**: e15098.
14. Lessler J, Cummings DA, Read JM, et al. Location-specific patterns of exposure to recent pre-pandemic strains of influenza A in southern China. *Nat Commun* 2011; **2**: 423.

11. Annexes – data capture forms

Data Capture Forms	Collected Data	Timeline
Annex 1 - Screening Interview	Initial participation willingness; reasons for rejection or ineligibility; basic information such as age, sex, self-rated health, awareness of influenza vaccine	Baseline
Annex 2 - Enrolment Interview 1	Contact information for follow up activities; demographic information; household information; health and mobility; blood samples	Baseline
Annex 3 - Enrolment Interview 2 (Version A & B)	Functional status, standardized mini-mental state examination (SMMSE), life history child to adult (2A only), life history young adult (2Aonly), life history, socioeconomic status, general health, chronic disease, and knowledge, attitudes, and practices (KAP) towards influenza vaccination	1 - 3 months after baseline
Annex 4 - Half Year Follow-up Interview	Mini survey of functional status; a sub-group of subjects were asked to provide blood samples	6 months after baseline
Annex 5 - Annual Reassessment	Reconfirmation of contact information; demographic information; household information; general health; chronic disease; functional status; mini vaccination KAP; SMMSE (Year 2 only); sera samples	Around the end of each year

Annex 6 - Medical Chart Review	Demographic information; past medical history	Ongoing
Annex 7 - Hospital Case Report	Demographic information; history of present illness; past medical history; vaccination history; treatment prior to hospitalization; hospital and admission information; clinical evaluation and vital signs at triage; treatment during hospitalization; testing results; discharge	Hospitalization due to acute episode since enrolment
Annex 8 - Symptom Screening Log	Symptoms; illness onset date	Upon identification of symptoms
Annex 9 - Acute Illness Interview	Illness background; symptoms and severity; Mini-Cog^{TM} tool; medical care; daily activities	Within 7 days since illness onset
Annex 10 - Illness Follow-up Interview	Symptoms and severity; illness resolution day; medical care; daily activities	Around the 10th day after illness onset
Annex 11 - Symptom Diary	A subset of participants will be asked to log the presence or absence of 12 symptoms and their highest temperature	Up to 10 days after swab collection
Annex 12 - Death Record	Confirmation of death; source of confirmation; date of death; cause of death	Upon confirmation of death

Annex 13 - Participant Withdrawal Form	Confirmation of withdrawal, type, date of withdrawal, reason for withdrawal	Upon confirmation of withdrawal
Annex 14 – Monthly Hospitalization Surveillance Form	Hospitalization information of participants	Last week of every month