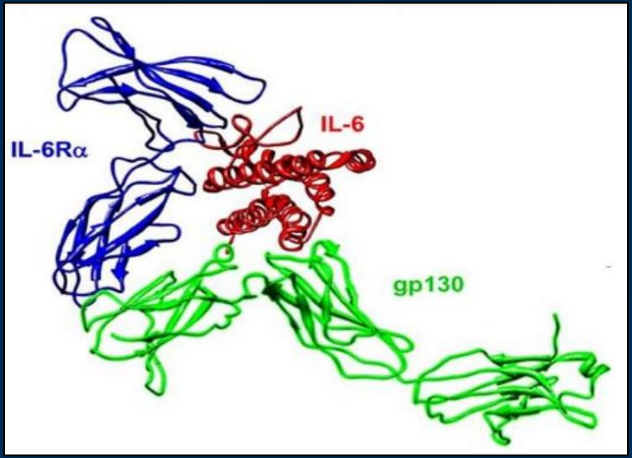
Seasonal variation in the association between the rs2228145 variant of the IL-6R gene and self-reported long-COVID symptoms: a clinical study



Katie Rees[1], Rebecca Aicheler[1], Lee Butcher[1], Alan Dodd[2], John Geen[2], Ceri Lynch[3], Isabel Massey[1], Keith Morris[1], Brian Tennant[2], Richard Webb[1]*

1 Department of Biomedical Sciences, School of Sport/Health Sciences, Cardiff Metropolitan University, Cardiff CF5 2YB, UK
2 Clinical Biochemistry Department, Cwm Taf Morgannwg University Health Board, Merthyr Tydfil CF47 9DT, UK
3 Intensive Care Unit, Cwm Taf Morgannwg University Health Board, Ynysmaerdy, Pontyclun CF72 8XR, UK

Cardiff Metropolitan University | Prifysgol Metropolitan Caerdydd

BUDDSODDWYR MEWN POBL | INVESTORS IN PEOPLE

Introduction [i]: long-COVID (PCS, PASC)

- **Definition: persistence for at least 12 weeks from the onset of COVID-19 of symptoms which cannot be explained by an alternative diagnosis** (WHO).
- Estimated to be approx 200 million affected individuals worldwide, and this total is projected to increase.
  - UK: ~1.7 million people in the experience long-COVID symptoms, for 59% of whom these symptoms impact on their day-to-day activities.

- WHO: **10-20% long-COVID prevalence**.
- [Hastie 2023],[O'Mahoney 2022],[Woodrow 2023]: **44-50% long-COVID prevalence.**

Q1: Why this discrepancy?

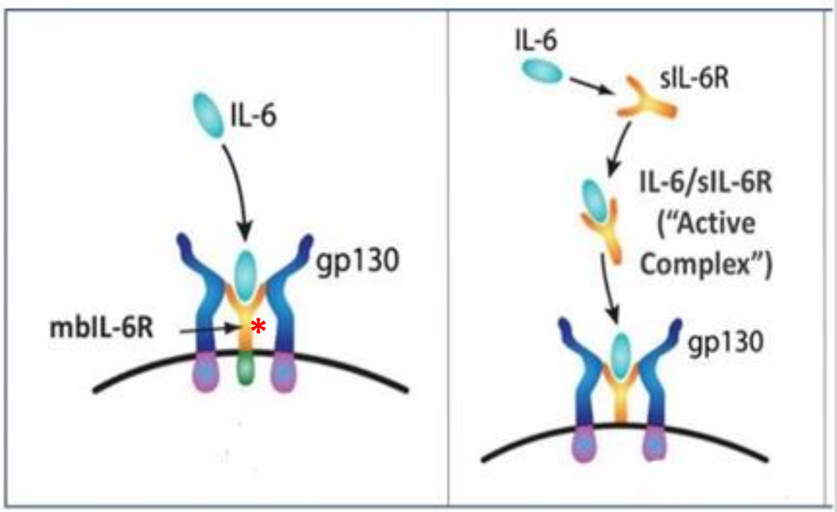
Q2: Is it possible to identify a marker that predicts which individuals are likely to experience long-COVID following a COVID-19 infection?

Cardiff Metropolitan University | Prifysgol Metropolitan Caerdydd

Introduction [ii]: IL-6/IL-6R/sgp130 signalling

- **rs2228145**: A→C change at position 1073 of **IL-6R** gene.

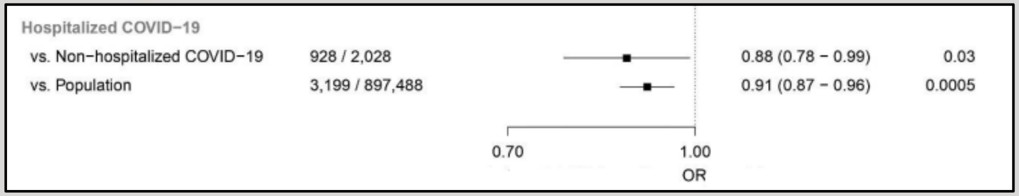
1073
DNA ...CAAG**A**TTCT...
G**A**T [Asp^{358}] → G**C**T [Ala^{358}]
Protein **...Gln Asp Ser...**
OR
...Gln Ala Ser...



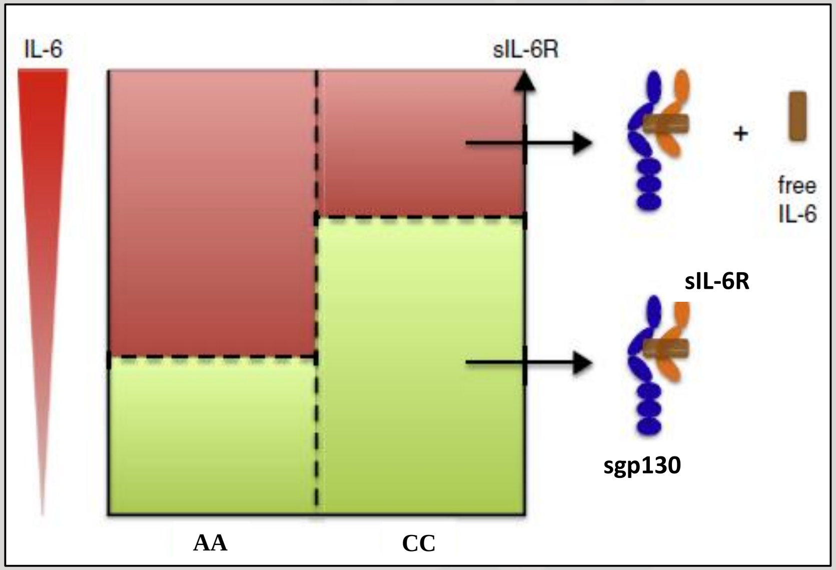
- **IL-6R protein** is cut at * into 2 fragments, one of which is shed from cells as ***sIL-6R***.
- Ala^{358} confers ↑ sensitivity to cleavage; **~2x [sIL-6R] in CC individuals** *(cf. AA or AC).*

> Garbers (2015): *"sIL-6R & sgp130 constitute a buffer for IL-6 in the blood. The rs2228145 CC genotype leads to higher blood-borne sIL-6R concentrations, & thus higher IL-6-buffering capacity. This leads to protection against ↑ unbuffered IL-6 in the context of inflammatory diseases."*

eg. Bovjin (2020): CC is **protective** re COVID-19 severity:



Hospitalized COVID−19			
vs. Non−hospitalized COVID−19	928 / 2,028	0.88 (0.78 − 0.99)	0.03
vs. Population	3,199 / 897,488	0.91 (0.87 − 0.96)	0.0005

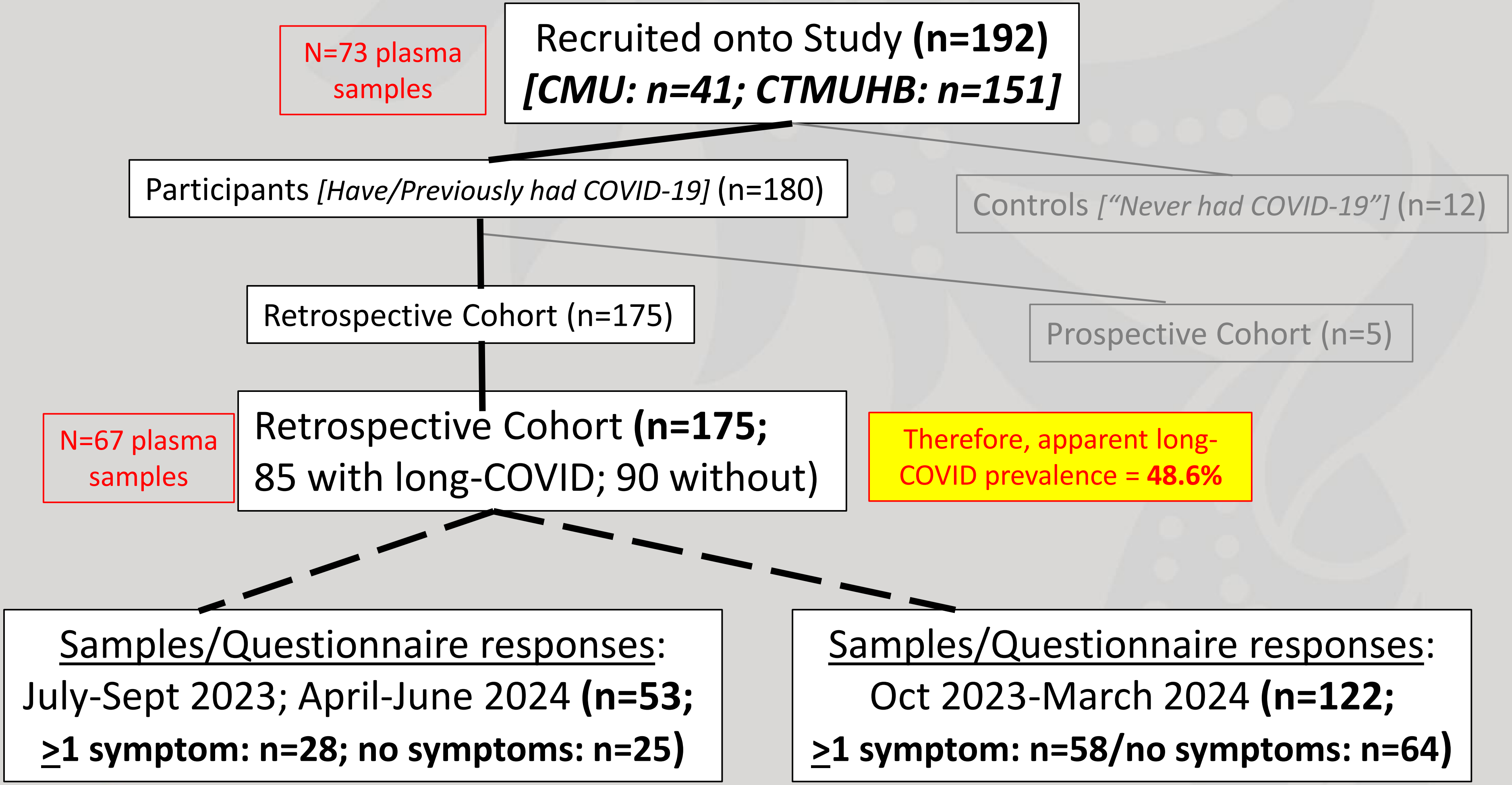


Study Aims/Objectives

- The study's aim is to **investigate the influence of IL-6 signalling** (particularly the impact of **the rs2228145 variant in the IL-6R gene**) on the **body's response to COVID** infection.
- **Primary Objective:** To obtain **DNA samples** from patients [i] with severe COVID-19, [ii] who have recovered from severe COVID-19, [iii] with long-COVID, and [iv] who have never had COVID-19, to determine their **rs2228145 genotype**.
- **Secondary Objectives**
  - a) To collect **plasma samples** & determine circulating levels of **IL-6, sIL-6R & sgp130**.
  - b) To evaluate potential associations between the above and: [i] **severity of patients' acute symptoms** during SARS-CoV-2 infection; [ii] **progression to long-COVID** in the medium-term aftermath of SARS-COV-2 infection.
  - c) To evaluate potential associations between long-COVID risk and factors such as **number/severity of COVID-19 infections, vaccination history, co-morbidities**.

Cardiff Metropolitan University | Prifysgol Metropolitan Caerdydd

Study Design



N=73 plasma samples

Recruited onto Study **(n=192)**
[CMU: n=41; CTMUHB: n=151]

Participants *[Have/Previously had COVID-19]* (n=180)

Controls *[“Never had COVID-19”]* (n=12)

Retrospective Cohort (n=175)

Prospective Cohort (n=5)

N=67 plasma samples

Retrospective Cohort **(n=175;**
85 with long-COVID; 90 without)

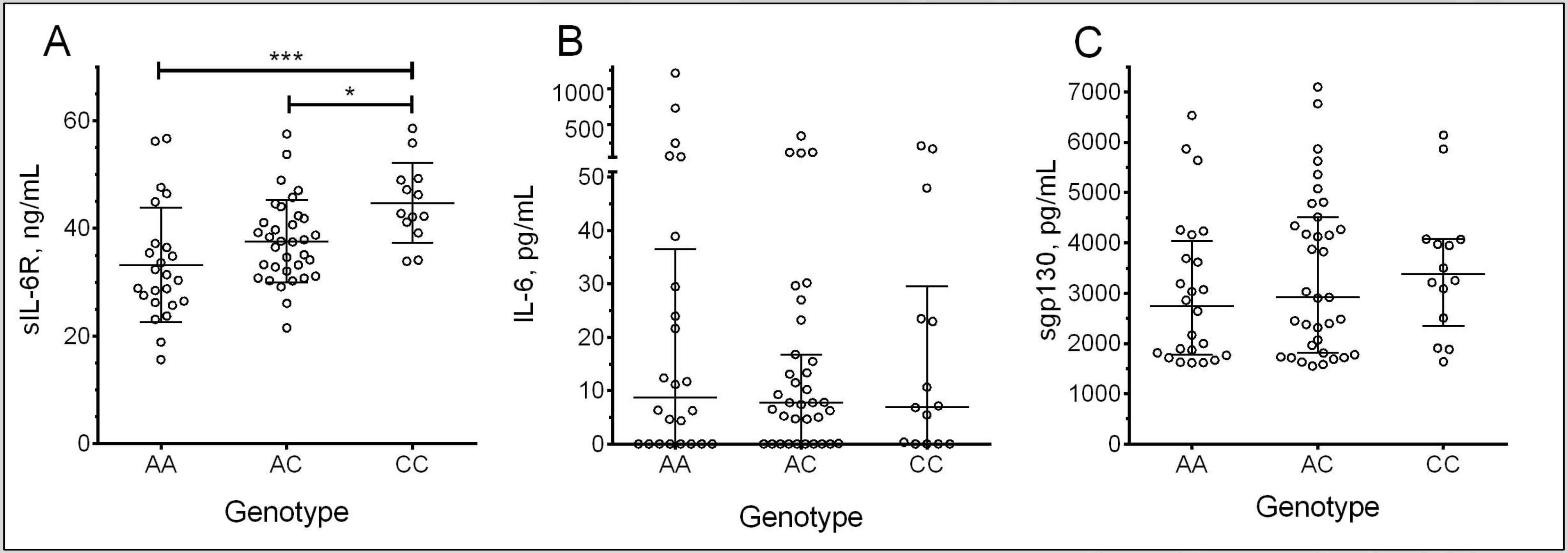
Therefore, apparent long-COVID prevalence = **48.6%**

Samples/Questionnaire responses:
July-Sept 2023; April-June 2024 **(n=53;**
≥1 symptom: n=28; no symptoms: n=25)

Samples/Questionnaire responses:
Oct 2023-March 2024 **(n=122;**
≥1 symptom: n=58/no symptoms: n=64)

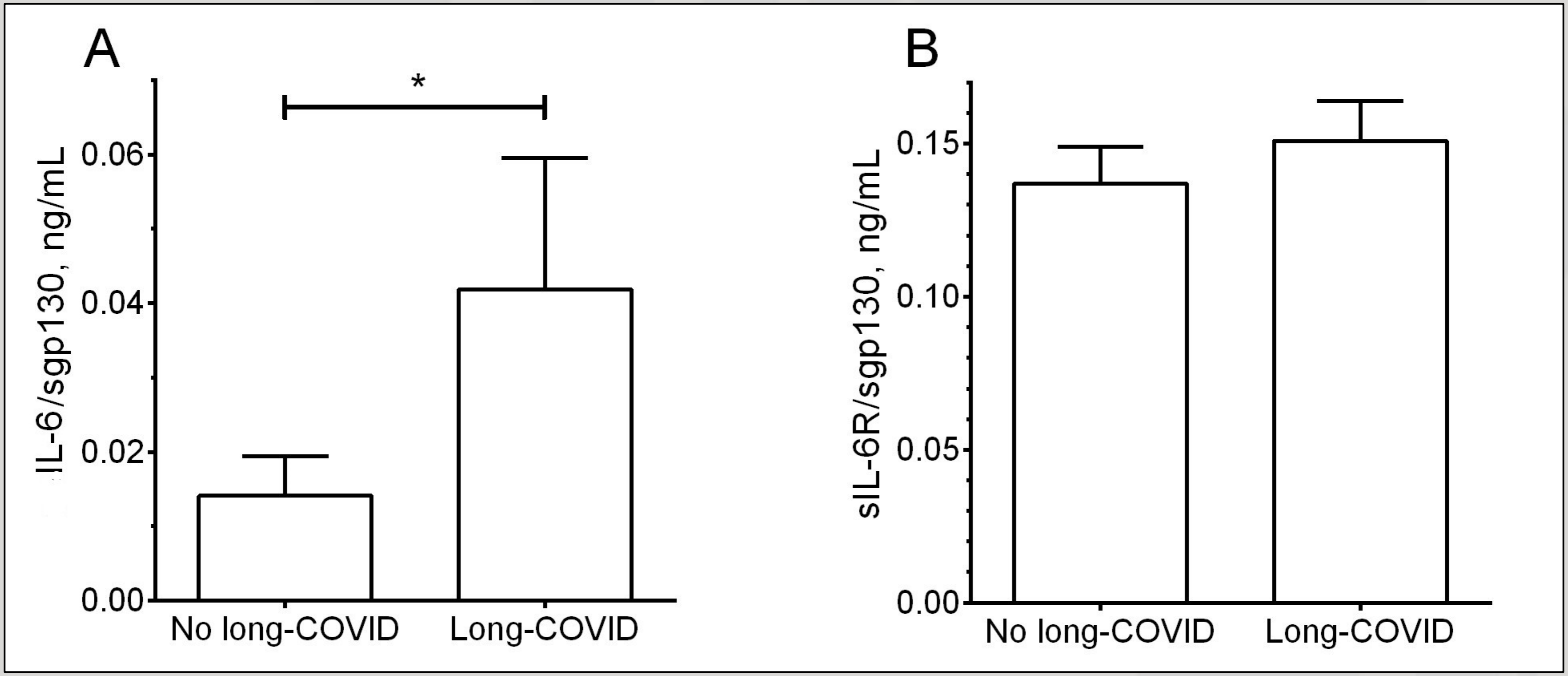
Cardiff Metropolitan University | Prifysgol Metropolitan Caerdydd

rs2228145 genotype significantly associated with sIL-6R levels (but not IL-6, sgp130)



Cardiff Metropolitan University | Prifysgol Metropolitan Caerdydd

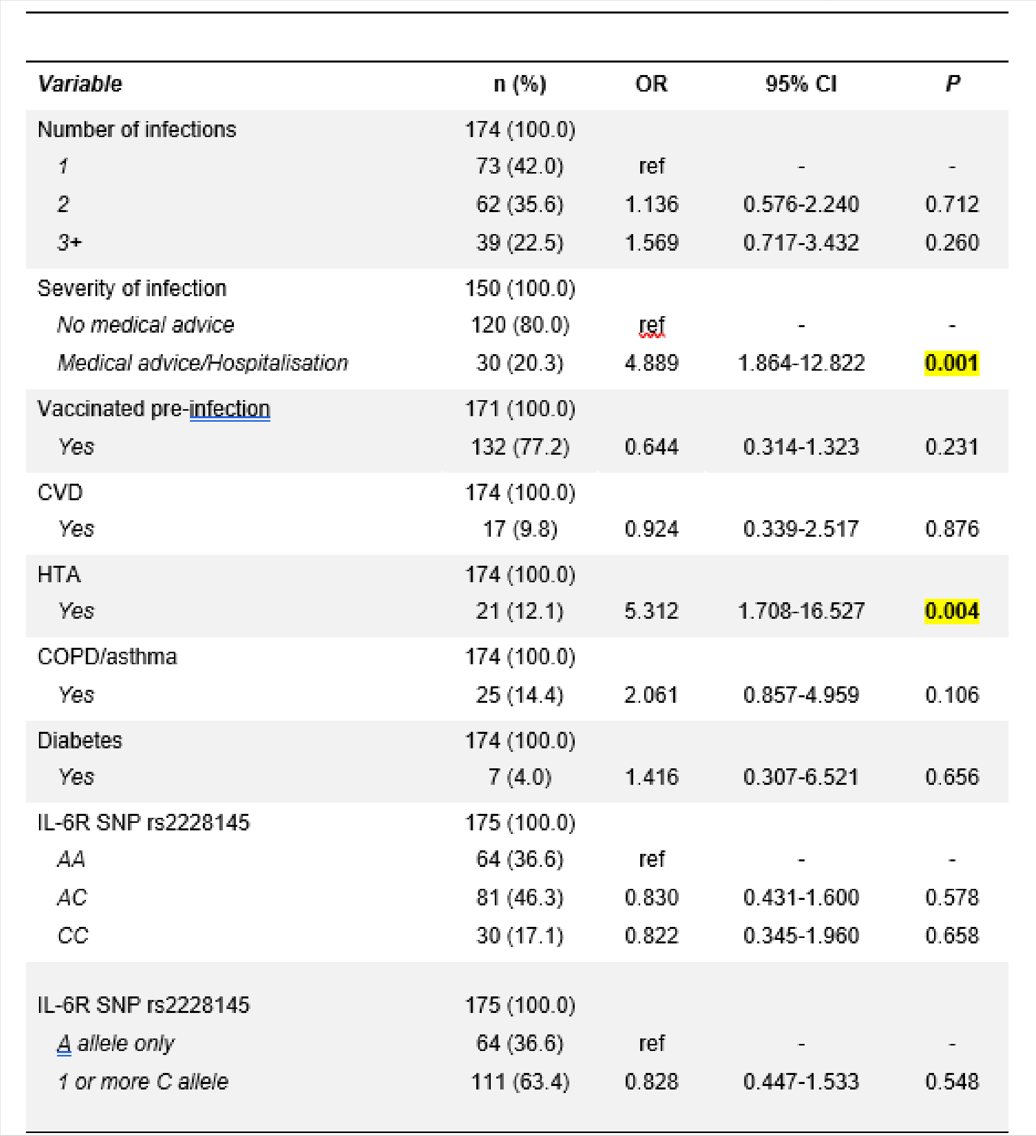
IL-6 levels significantly associated with presence of Long-COVID



- Absolute IL-6 values: **78±36pg/ml [long-COVID] versus 29±13pg/ml (healthy)**.
  - *Guirao (2020) "A cut-off point of **35pg/mL** clearly differentiates patients with more severe disease"*

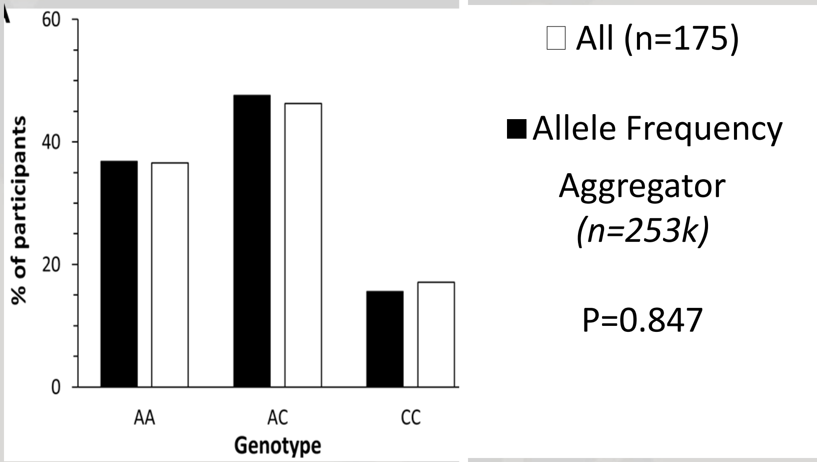
Cardiff Metropolitan University | Prifysgol Metropolitan Caerdydd

Univariate logistic regression analysis for potential contributing factors of long-covid *(1 or more symptoms)*

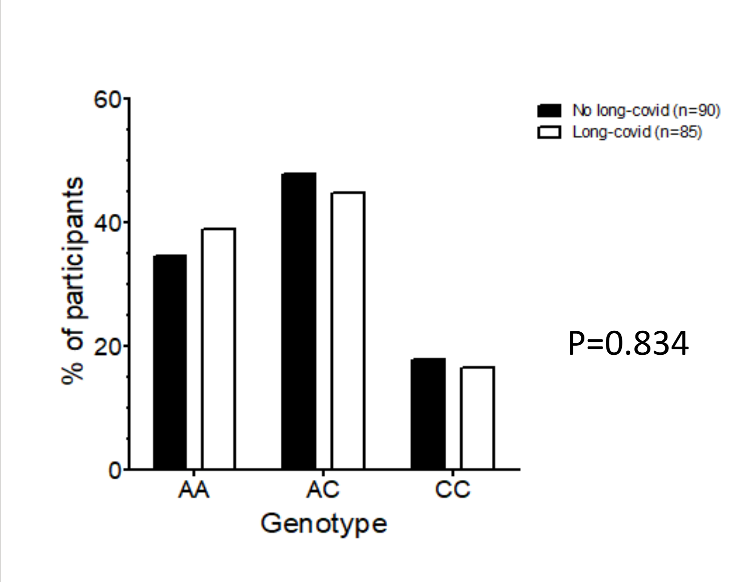


Variable	n (%)	OR	95% CI	*P*
Number of infections	174 (100.0)			
1	73 (42.0)	ref	-	-
2	62 (35.6)	1.136	0.576-2.240	0.712
3+	39 (22.5)	1.569	0.717-3.432	0.260
Severity of infection	150 (100.0)			
No medical advice	120 (80.0)	ref	-	-
Medical advice/Hospitalisation	30 (20.3)	4.889	1.864-12.822	0.001
Vaccinated pre-infection	171 (100.0)			
Yes	132 (77.2)	0.644	0.314-1.323	0.231
CVD	174 (100.0)			
Yes	17 (9.8)	0.924	0.339-2.517	0.876
HTA	174 (100.0)			
Yes	21 (12.1)	5.312	1.708-16.527	0.004
COPD/asthma	174 (100.0)			
Yes	25 (14.4)	2.061	0.857-4.959	0.106
Diabetes	174 (100.0)			
Yes	7 (4.0)	1.416	0.307-6.521	0.656
IL-6R SNP rs2228145	175 (100.0)			
AA	64 (36.6)	ref	-	-
AC	81 (46.3)	0.830	0.431-1.600	0.578
CC	30 (17.1)	0.822	0.345-1.960	0.658
IL-6R SNP rs2228145	175 (100.0)			
A allele only	64 (36.6)	ref	-	-
1 or more C allele	111 (63.4)	0.828	0.447-1.533	0.548

Genotype Frequency



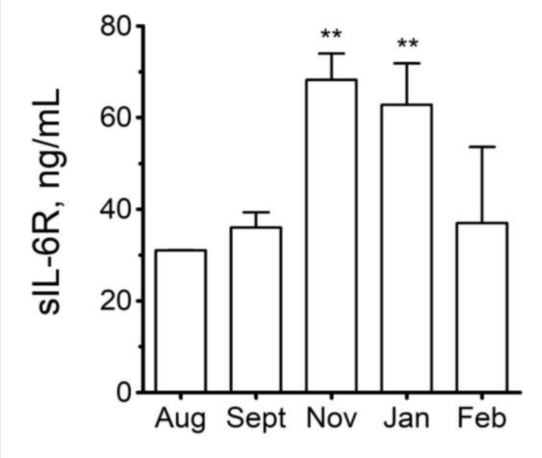
P=0.847



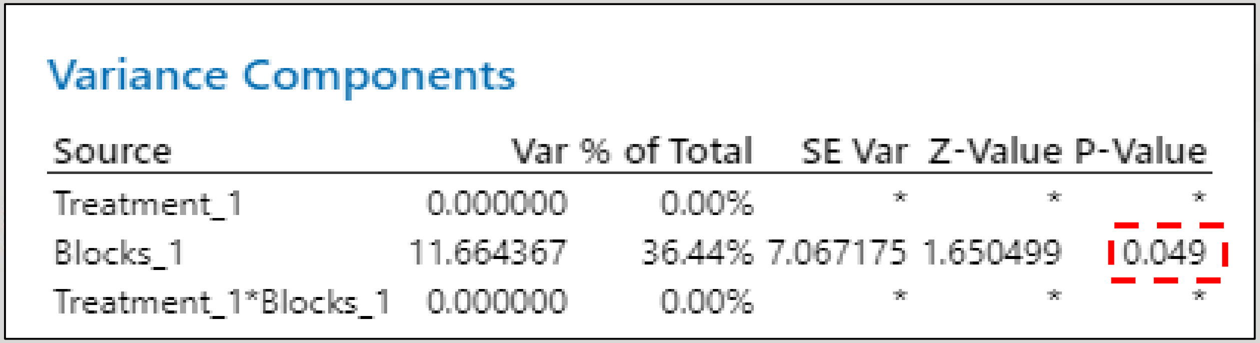
P=0.834

Stratification, on the basis of time of year

- **Prev. studies: COVID-19, IL-6 & sIL-6R are significantly affected by season.**
  - Maes (1994)
  - De Vita (2014)
  - **Webb (2021)** Provisional Designation of IL-6R as a Novel Marker Gene for Exercise Tolerance
  - Byun (2021). "*Is coronavirus disease (COVID-19) seasonal? A critical analysis*"
  - Walsh (2023)



- **Mixed-effects model analysis:** Long-COVID score is significantly affected by season in our dataset ($P<0.05$).
  - Season ('*block*') & genotype ('*treatment*') v. Long-COVID score

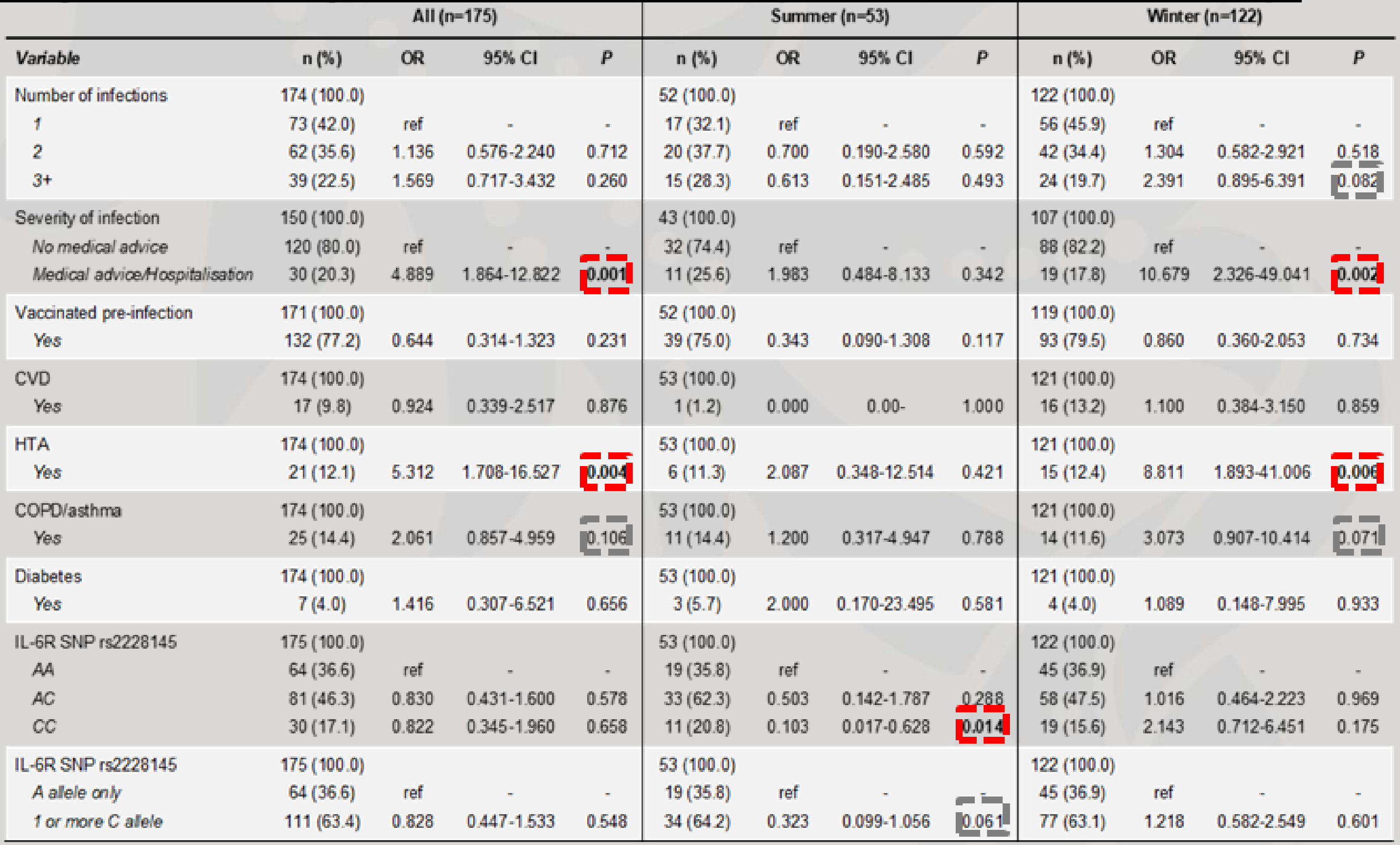


Variance Components

Source	Var	% of Total	SE Var	Z-Value	P-Value
Treatment_1	0.000000	0.00%	*	*	*
Blocks_1	11.664367	36.44%	7.067175	1.650499	0.049
Treatment_1*Blocks_1	0.000000	0.00%	*	*	*

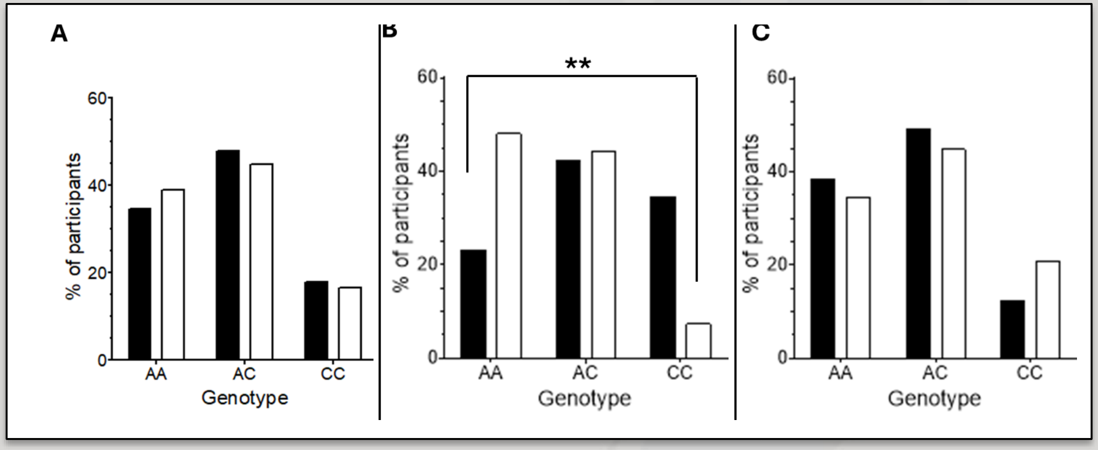
Therefore, dataset stratified into 'summer' & 'winter' sub-cohorts

Univariate logistic regression analysis *(seasonally stratified)*

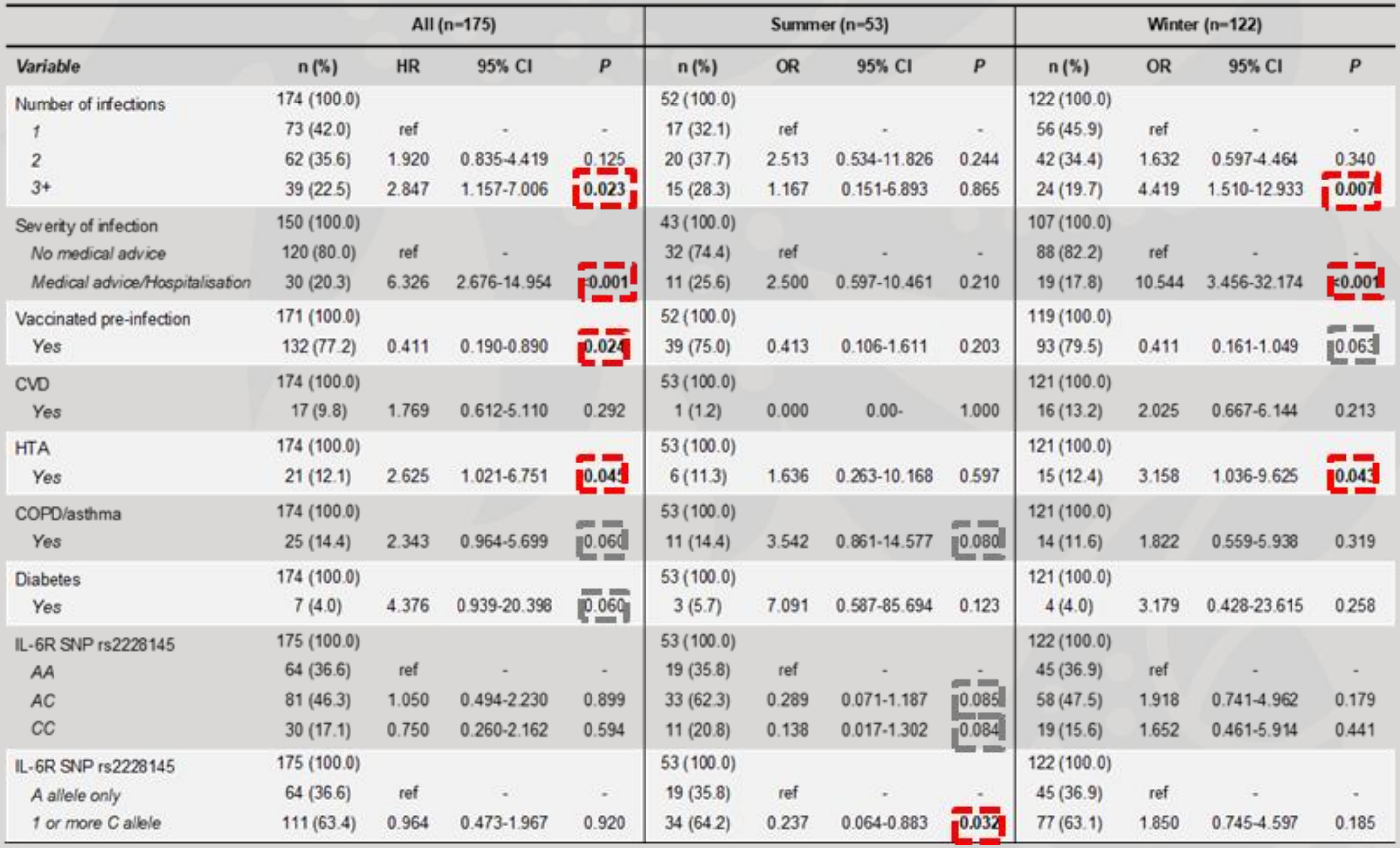


	All (n=175)				Summer (n=53)				Winter (n=122)			
Variable	n (%)	OR	95% CI	*P*	n (%)	OR	95% CI	*P*	n (%)	OR	95% CI	*P*
Number of infections	174 (100.0)				52 (100.0)				122 (100.0)			
1	73 (42.0)	ref	-	-	17 (32.1)	ref	-	-	56 (45.9)	ref	-	-
2	62 (35.6)	1.136	0.576-2.240	0.712	20 (37.7)	0.700	0.190-2.580	0.592	42 (34.4)	1.304	0.582-2.921	0.518
3+	39 (22.5)	1.569	0.717-3.432	0.260	15 (28.3)	0.613	0.151-2.485	0.493	24 (19.7)	2.391	0.895-6.391	0.082
Severity of infection	150 (100.0)				43 (100.0)				107 (100.0)			
No medical advice	120 (80.0)	ref	-	-	32 (74.4)	ref	-	-	88 (82.2)	ref	-	-
Medical advice/Hospitalisation	30 (20.3)	4.889	1.864-12.822	0.001	11 (25.6)	1.983	0.484-8.133	0.342	19 (17.8)	10.679	2.326-49.041	0.002
Vaccinated pre-infection	171 (100.0)				52 (100.0)				119 (100.0)			
Yes	132 (77.2)	0.644	0.314-1.323	0.231	39 (75.0)	0.343	0.090-1.308	0.117	93 (79.5)	0.860	0.360-2.053	0.734
CVD	174 (100.0)				53 (100.0)				121 (100.0)			
Yes	17 (9.8)	0.924	0.339-2.517	0.876	1 (1.2)	0.000	0.00-	1.000	16 (13.2)	1.100	0.384-3.150	0.859
HTA	174 (100.0)				53 (100.0)				121 (100.0)			
Yes	21 (12.1)	5.312	1.708-16.527	0.004	6 (11.3)	2.087	0.348-12.514	0.421	15 (12.4)	8.811	1.893-41.006	0.006
COPD/asthma	174 (100.0)				53 (100.0)				121 (100.0)			
Yes	25 (14.4)	2.061	0.857-4.959	0.106	11 (14.4)	1.200	0.317-4.947	0.788	14 (11.6)	3.073	0.907-10.414	0.071
Diabetes	174 (100.0)				53 (100.0)				121 (100.0)			
Yes	7 (4.0)	1.416	0.307-6.521	0.656	3 (5.7)	2.000	0.170-23.495	0.581	4 (4.0)	1.089	0.148-7.995	0.933
IL-6R SNP rs2228145	175 (100.0)				53 (100.0)				122 (100.0)			
AA	64 (36.6)	ref	-	-	19 (35.8)	ref	-	-	45 (36.9)	ref	-	-
AC	81 (46.3)	0.830	0.431-1.600	0.578	33 (62.3)	0.503	0.142-1.787	0.288	58 (47.5)	1.016	0.464-2.223	0.969
CC	30 (17.1)	0.822	0.345-1.960	0.658	11 (20.8)	0.103	0.017-0.628	0.014	19 (15.6)	2.143	0.712-6.451	0.175
IL-6R SNP rs2228145	175 (100.0)				53 (100.0)				122 (100.0)			
A allele only	64 (36.6)	ref	-	-	19 (35.8)	ref	-	-	45 (36.9)	ref	-	-
1 or more C allele	111 (63.4)	0.828	0.447-1.533	0.548	34 (64.2)	0.323	0.099-1.056	0.061	77 (63.1)	1.218	0.582-2.549	0.601

rs2228145 genotype frequency



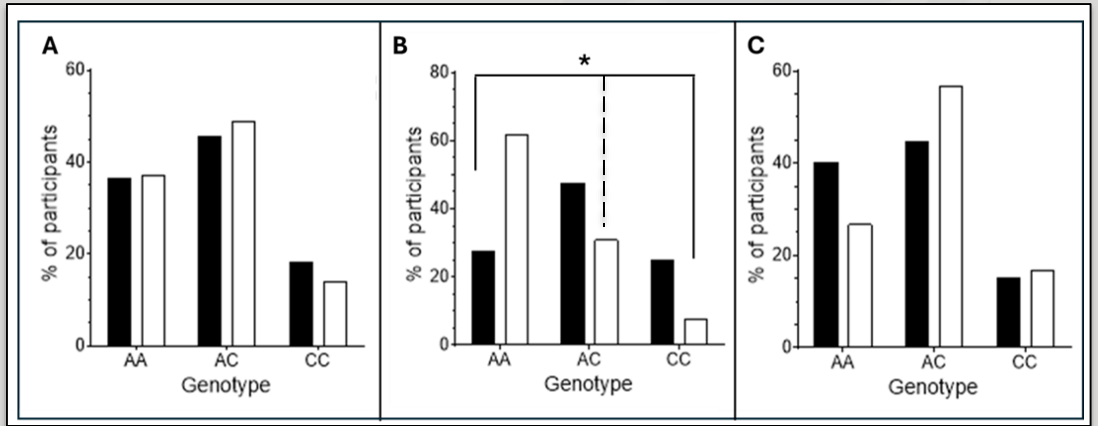
Univariate logistic regression analysis *(chronic fatigue)*



	All (n=175)				Summer (n=53)				Winter (n=122)			
Variable	n (%)	HR	95% CI	*P*	n (%)	OR	95% CI	*P*	n (%)	OR	95% CI	*P*
Number of infections	174 (100.0)				52 (100.0)				122 (100.0)			
1	73 (42.0)	ref	-	-	17 (32.1)	ref	-	-	56 (45.9)	ref	-	-
2	62 (35.6)	1.920	0.835-4.419	0.125	20 (37.7)	2.513	0.534-11.826	0.244	42 (34.4)	1.632	0.597-4.464	0.340
3+	39 (22.5)	2.847	1.157-7.006	0.023	15 (28.3)	1.167	0.151-6.893	0.865	24 (19.7)	4.419	1.510-12.933	0.007
Severity of infection	150 (100.0)				43 (100.0)				107 (100.0)			
No medical advice	120 (80.0)	ref	-		32 (74.4)	ref	-	-	88 (82.2)	ref	-	-
Medical advice/Hospitalisation	30 (20.3)	6.326	2.676-14.954	<0.001	11 (25.6)	2.500	0.597-10.461	0.210	19 (17.8)	10.544	3.456-32.174	<0.001
Vaccinated pre-infection	171 (100.0)				52 (100.0)				119 (100.0)			
Yes	132 (77.2)	0.411	0.190-0.890	0.024	39 (75.0)	0.413	0.106-1.611	0.203	93 (79.5)	0.411	0.161-1.049	0.063
CVD	174 (100.0)				53 (100.0)				121 (100.0)			
Yes	17 (9.8)	1.769	0.612-5.110	0.292	1 (1.2)	0.000	0.00-	1.000	16 (13.2)	2.025	0.667-6.144	0.213
HTA	174 (100.0)				53 (100.0)				121 (100.0)			
Yes	21 (12.1)	2.625	1.021-6.751	0.045	6 (11.3)	1.636	0.263-10.168	0.597	15 (12.4)	3.158	1.036-9.625	0.043
COPD/asthma	174 (100.0)				53 (100.0)				121 (100.0)			
Yes	25 (14.4)	2.343	0.964-5.699	0.060	11 (14.4)	3.542	0.861-14.577	0.080	14 (11.6)	1.822	0.559-5.938	0.319
Diabetes	174 (100.0)				53 (100.0)				121 (100.0)			
Yes	7 (4.0)	4.376	0.939-20.398	0.060	3 (5.7)	7.091	0.587-85.694	0.123	4 (4.0)	3.179	0.428-23.615	0.258
IL-6R SNP rs2228145	175 (100.0)				53 (100.0)				122 (100.0)			
AA	64 (36.6)	ref	-	-	19 (35.8)	ref	-	-	45 (36.9)	ref	-	-
AC	81 (46.3)	1.050	0.494-2.230	0.899	33 (62.3)	0.289	0.071-1.187	0.085	58 (47.5)	1.918	0.741-4.962	0.179
CC	30 (17.1)	0.750	0.260-2.162	0.594	11 (20.8)	0.138	0.017-1.302	0.084	19 (15.6)	1.652	0.461-5.914	0.441
IL-6R SNP rs2228145	175 (100.0)				53 (100.0)				122 (100.0)			
A allele only	64 (36.6)	ref	-	-	19 (35.8)	ref	-	-	45 (36.9)	ref	-	-
1 or more C allele	111 (63.4)	0.964	0.473-1.967	0.920	34 (64.2)	0.237	0.064-0.883	0.032	77 (63.1)	1.850	0.745-4.597	0.185

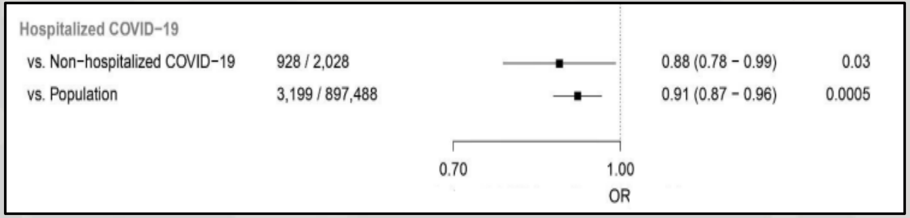
rs2228145 genotype frequency

■ No long-covid
□ Long-covid



Discussion.

- **‘Summer cohort’ data**: **significant protective effect of CC in context of long-COVID**; aligns with prev. COVID-19 studies: *eg. Bovjin (2020).*



  - Adds a **novel ‘long-COVID’ context** to these previous studies.

- **Seasonal Misclassification Bias in self-reported symptom studies**:
  - **Hastie, 2023:** 64.5% of ’long-COVID’ cohort reported apparent long-COVID symptoms, but so did 50.8% of the ‘never-had-COVID’ cohort! ***Adjusting for bias → reduction of initial estimated prevalence of ~50% to 13.8%*** *(in line with WHO estimate).*
  - Long-COVID-resembling symptoms have been reported for several seasonal conditions *(eg. Seasonal chronic fatigue syndrome, seasonal allergy-related conditions, seasonal non-COVID respiratory diseases, seasonal depression).*
  - ‘Capture’ of long-COVID via questionnaire responses *(ie. self-reported symptoms)* may be less accurate in winter, due to the confounding effect of the higher prevalence of seasonal long-COVID-resembling non-COVID conditions during this period.

Cardiff Metropolitan University | Prifysgol Metropolitan Caerdydd

Conclusion

- While interpretation is complicated by seasonal variation, the finding that **rs2228145 CC genotype may be linked to protection from developing long-COVID may be of clinical value**.
- Possible future use of rs2228145 genotype as a **biomarker** predictive of long-COVID risk? *May be of use in identifying at-risk individuals, facilitating rapid signposting towards relevant treatment options, and bringing advantages regarding long-COVID management/treatment.*

Acknowledgements

- **Cardiff Met**: Katie Rees, Is Massey, Prof Keith Morris, Dr Becky Aicheler, Dr Lee Butcher
  - *Also: Dan Nash, Saira Sarwar, Alison Early, Dr Mike Hughes, Dr Rachel Sumner*
- **CTMUHB**: Prof John Geen, Dr Ceri Lynch, Dr Brian Tennant, Dr Alan Dodd, Katy-May Price, Sarah Gill, Alysha Hancock, Samantha Evans, Meryl Rees, Lauren Geen, Lisa Roche, Keri Turner
- Financial support: **CTMUHB Collaborative Research Fund** (+ Cardiff Met internal match-funding)

Cardiff Metropolitan University | Prifysgol Metropolitan Caerdydd

GIG CYMRU NHS WALES | Bwrdd Iechyd Prifysgol Cwm Taf Morgannwg University Health Board