TITLE : CHARACTERISTICS OF PARTICIPANTS OBTAINED FROM THE DATA VISUALIZATION TOOLS OF THE *BIOBANQUE QUÉBÉCOISE DE LA COVID-19*; A POWERFUL OPEN SCIENCE TOOL TO ADVANCE COVID-19 RESEARCH.

Madeleine Durand, David Bujold, Guillaume Bourque, Karine Tremblay, Emilia Liana Falcone, Alain Piché, Simon Décary, Sze Man Tse, Christel Renoux, Guillaume Bourque, Ma'n Zawati, Daniel Auld, Alexandre Monpetit, David Bujold, Brent Richards, Mylène Bertrand, Solomia Yanishevsky, Doris Ramsay, Pascale Léon, Marc Messier-Peet, Daniel Kaufmann, Vincent Mooser and Simon Rousseau, *for the Biobanque Québécoise de la COVID-19.*

Oral presentation at the CITF confenrence, Vancouver, Canada, March 8-9th

COI

- I have not conflict of interest to disclose

- BQC 19 is funded by FRQS, PHAC, Ministère de la sasnté et des services sociaux du Québec, Génome Québec, CITF

OBJECTIVES

- Understand the dynamics underlying participation to BQC 19

- Overview of characteristics of participants

- Familiarize ourselves with type of available content and data

- Discover the Bento data visualization platform

Participants (15 février 2023)
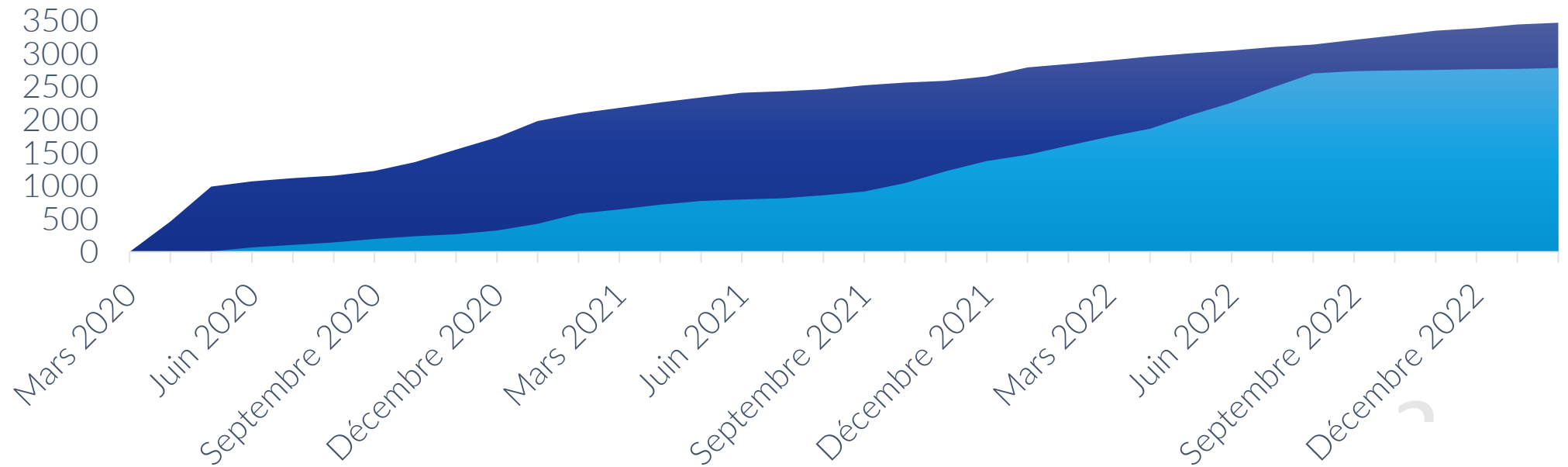
Participants	Participants 5e vague	Participants	Participants réinfectés
*Financement 2020-2022	* Recrutés entre 26 déc 2021 et 31 mars 2022	*Financement 2022-23	
4 037	603	1 612	398

Visites (15 février 2023)
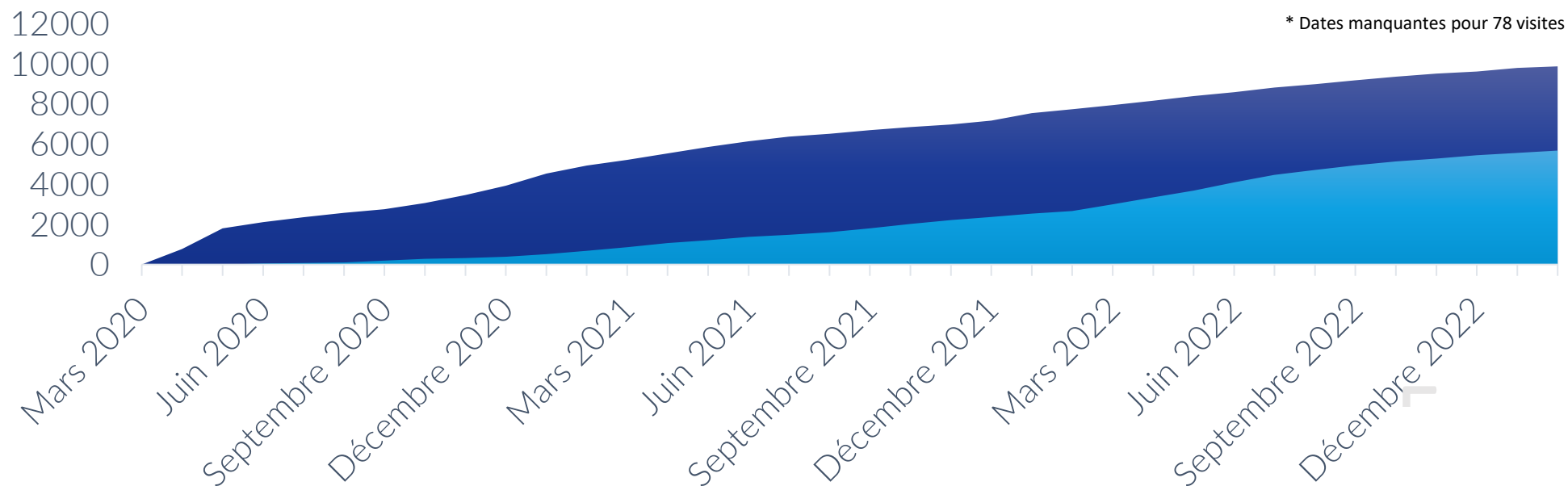
* Dates manquantes pour 78 visites

Visites	Visites 5e vague	Visites	Visites réinfection
*Financement 2020-2022	*Recrutés entre 26 déc 2021 et 31 mars 2022	*Financement 2022-23	
11 828	1 479	2 335	304

Overlap of multi*omics analysis per patients
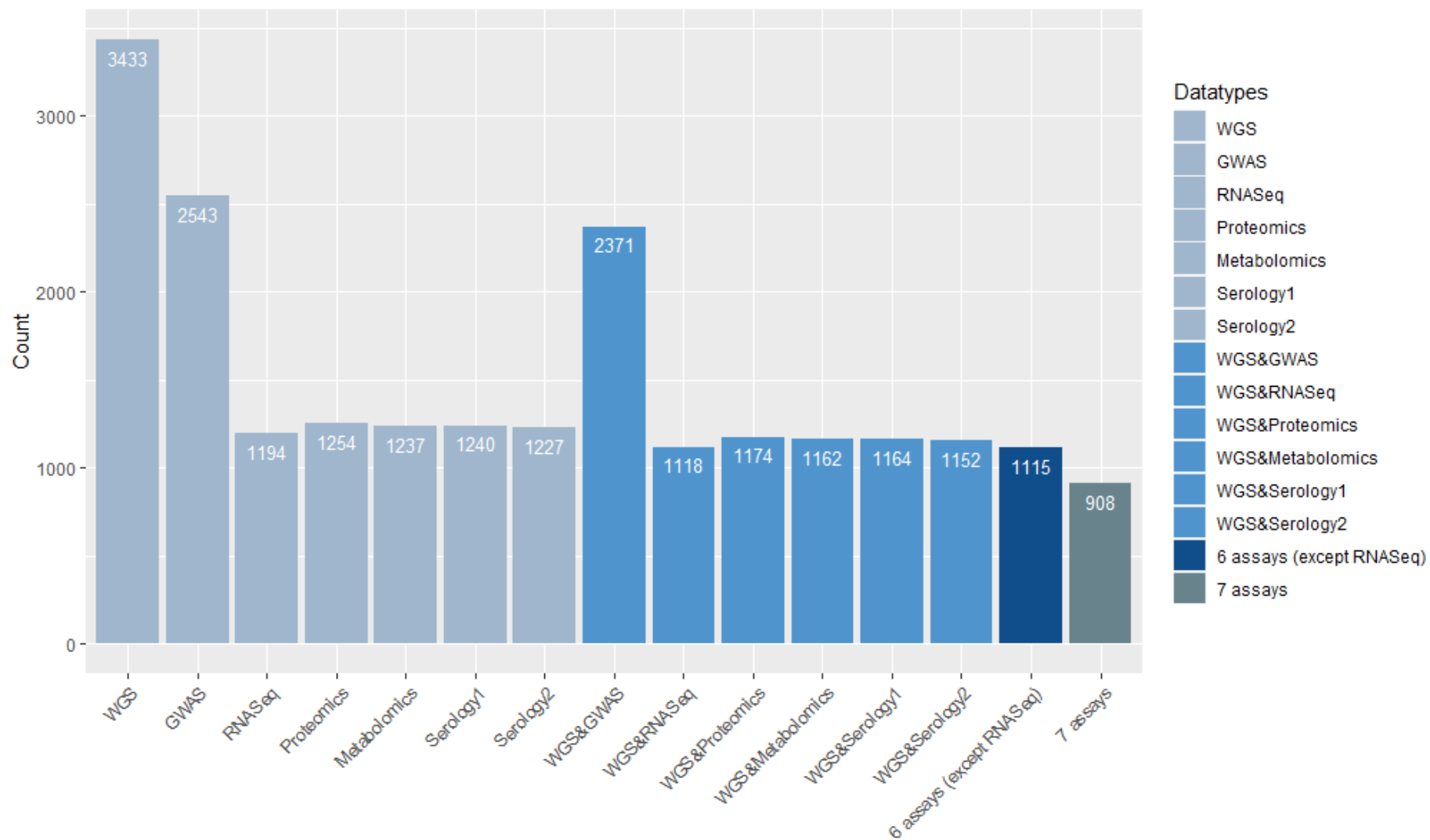
BASELINE CHARACTERISTICS

	Inpatients n=2581	Outpatient n=3074
Age (years, mean, SD)	59.85 (21.8)	43.81 (14.81)
Age categories		
≤17	157 (6%)	39 (1%)
18-50	723 (28%)	1708 (56%)
51-64	815 (32%)	618 (20%)
65-79	893 (35%)	205 (7%)
80+	545 (21%)	18 (0.6%)
Female sex (n, %)	1420 (55%)	1790 (58%)
Admission modality (n, %)		
ICU	651 (25.2%)	
Covid status (n, %)		
Positive	1814 (70%)	2943 (96%)
Negative	767 (30%)	131 ((4%)

	inpatients	outpatients
Comorbidities (n, %)	N=2581	N=3074
Hypertension	1318 (51%)	300 (10%)
Asthma	325 (13%)	259 (8%)
Chronic obstructive pulmonary disease	304 (12%)	37 (1%)
Diabetes	763 (30%)	124 (4%)
Stroke	160 (6%)	13 (0.4%)
Heart failure	249 (10%)	6 (0.2%)
Myocardial infarction	170 (6%)	13 (0.4%)
Chronic kidney disease stages≥3	343 (13%)	24 (0.8%)
Clinical frailty scale (n, %)	N=668	N=65
Non-frail (1-4)	505 (76%)	64 (98%)
Mild-to-moderately frail (5-6)	122 (18%)	1 (2%)
Severely frail (7+)	41 (6%)	0 (0%)

VACCINATION STATUS

| Vaccination status	
At least 1 dose	3030
At least 2 doses	2682
At least 3 doses	1602
At least 4 doses	215

Etc, etc, etc...

BIOBANK'S CHALLENGES IN CONTENT DESCRIPTION
(AS A BARRIER FOR THEIR USE AND FAIR PRINCIPLES)

- Huge quantities of biosamples and associated data that are difficult to describe
  - Different populations
  - Long timelines
  - Variety of research questions and perspectives
  - Nature of biobank is to have a living inventory (*so what is left in it now?*)

- How can we assist researchers in optimizing biobank's use?

BENTO'S DATA VISUALIZATION PLATFORM
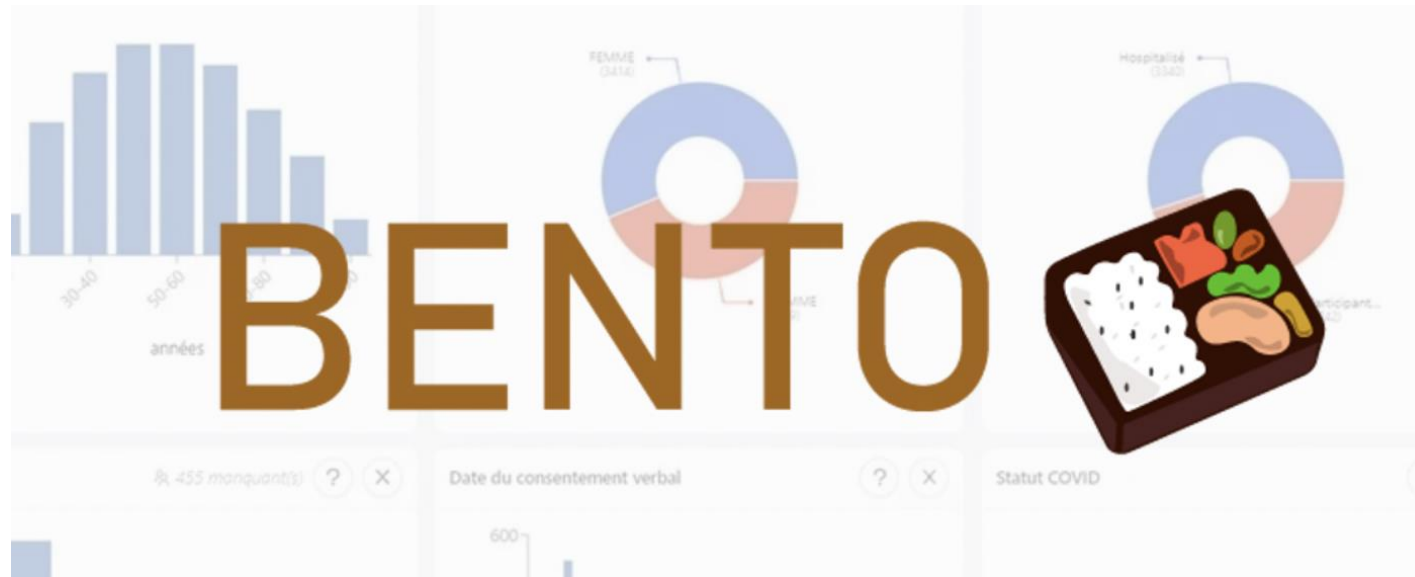
OK... BUT IS THERE WHAT I WOULD NEED IN BQC-19?

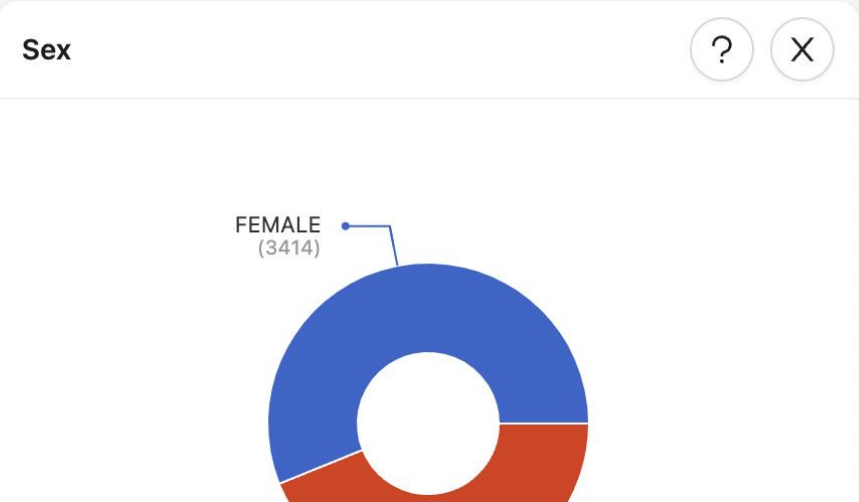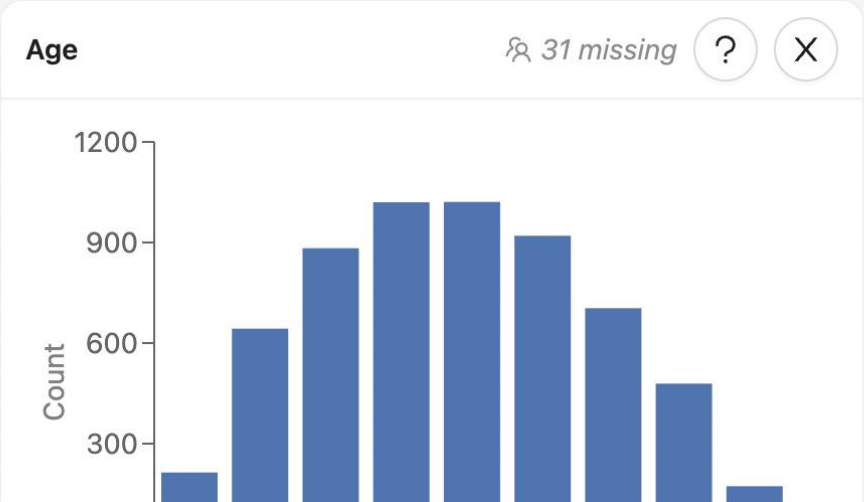https://en.quebeccovidbiobank.ca/bento

BQC19 Bento

Overview   **Search**   **Provenance**
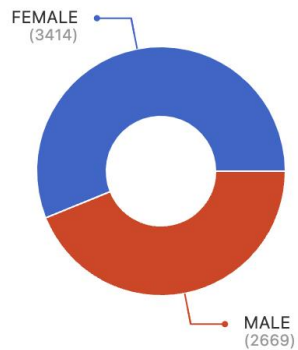
Individuals: 6088

La plateforme Bento permet à la communauté scientifique d'explorer les données agrégées de la cohorte BQC19. L'objectif est de permettre aux utilisateurs potentiels de la BQC19 de déterminer si les données disponibles dans la biobanque répondent aux besoins de leur projet de recherche avant qu'ils ne préparent une demande d'accès. Les utilisateurs dont les demandes d'accès ont été acceptées pourront bientôt accéder à des informations granulaires grâce à un portail privé sécurisé accessible par mot de passe.

The Bento platform enables the research community to explore the BQC19 cohort aggregate data. It is meant for potential data users to help them determine whether the cohort would be a good match for their research needs, before going through an access application process. Data and sample applicants will soon be allowed to access the granular information in a private, password-protected part of the portal, after obtaining proper access.
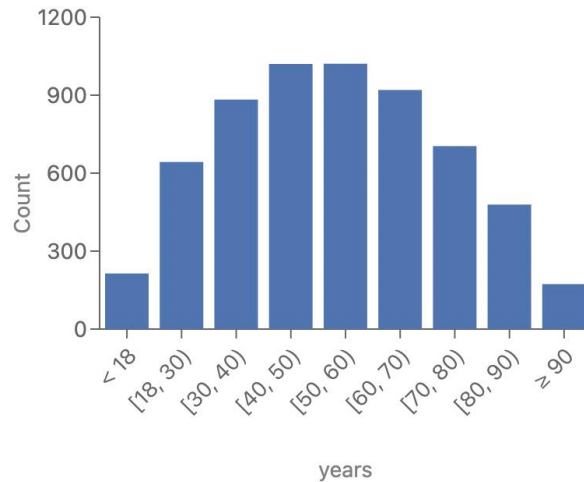
General

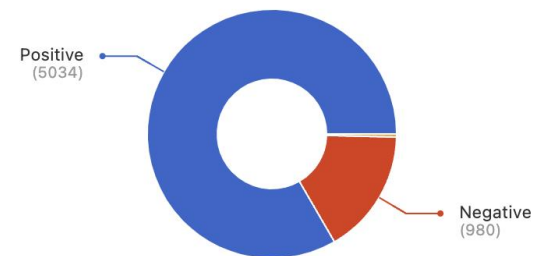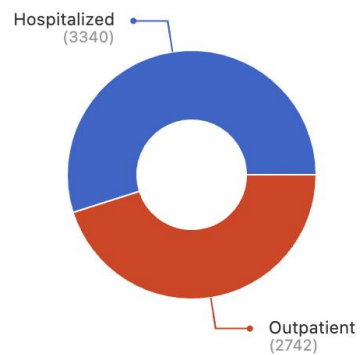**Age**   👥 *31 missing*   ?   ✕

1200

900

600

Count

300

Sex   ?   ✕

FEMALE
(3414)

Sex

FEMALE (3414)
MALE (2669)

Age

31 missing

Age histogram with Count (y-axis) up to 1200 and years (x-axis): < 18, [18, 30), [30, 40), [40, 50), [50, 60), [60, 70), [70, 80), [80, 90), ≥ 90

COVID status

Positive (5034)
Negative (980)

Hospitalization

6 missing

Hospitalized (3340)
Outpatient (2742)

Smoking

455 missing

Smoking bar chart with Count (y-axis) up to 3600: Non-smoker, Smoker, Former smoker, Passive smoker, Not specified
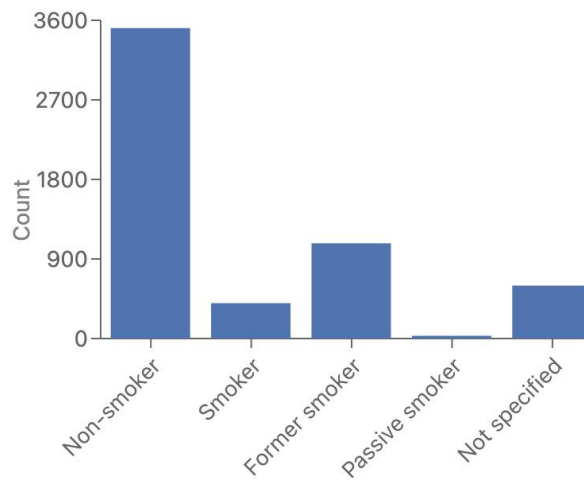
Diabetes

529 missing

Yes (946)
No (4612)
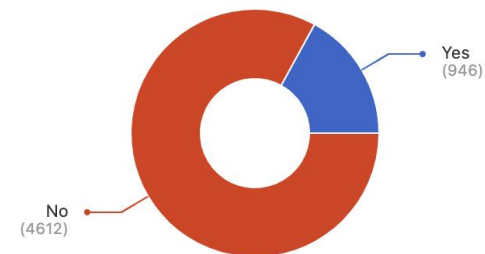
Manage Charts

General

- ☑ Age
- ☑ Sex
- ☐ Hospitalization
- ☐ Smoking
- ☐ Verbal consent date
- ☑ COVID status
- ☐ Vital status
- ☐ BMI
- ☑ ICU
- ☐ Participant category
- ☐ Domicile
- ☑ Vaccinal status
- ☑ Vaccine dose

Comorbidities

- ☐ Mobility
- ☐ Self-care
- ☑ Cardiac history
- ☐ TIA
- ☑ Asthma
- ☑ Stroke
- ☑ Diabetes
- ☑ Cancer
- ☑ Dementia
- ☐ Atrial fibrillation
- ☐ HIV
- ☑ Arterial Hypertension
- ☐ Immunosupression
- ☐ Myocardial infarction
- ☐ Heart failure
- ☑ Coronary artery disease
- ☑ Liver disease
- ☑ COPD
- ☑ Psychiatric disease
- ☑ Chronic kidney disease

Lab results

- ☑ CRP
- ☑ Creatinine
- ☐ Chest X-ray
- ☑ WBC
- ☑ Haemoglobin
- ☐ NT-proBNP
- ☑ LDH
- ☑ D-Dimer

Treatments and outcomes

- ☑ Respiratory support
- ☑ Reinfection
- ☑ Persisting symptoms
- ☐ Self-care post covid
- ☑ Systemic corticosteroid
- ☑ Antocoagulants
- ☐ ARDS
- ☐ Deep vein thrombosis
- ☐ Pulmonary embolism
- ☐ Acute kidney injury
- ☐ Stroke

BQC19 Bento

Overview **Search** Provenance

Individuals

Biosamples

Experiments

Biosamples

No Results

Experiments

No Results

General

Age ? < 18

Sex ? FEMALE

Get Data

en Portal

Individuals

👥 204

Biosamples

🧬 1,235

Experiments

⚗ 330

Biosamples

PBMC
(264)

Plasma ACD
(292)

Blood ACD
Metabolite ... (150)
(35)

Proteomic p...
(35)

Serum
(215)

RNA extract
(277)

Experiments

Genotyping
(62)

Enzyme-link...
(35)

Antibody me...
(34)

WGS
(99)

RNA-Seq
(30)

General

	Age	⦵	< 18 ⌄
☑	Sex	⦵	FEMALE ⌄
	Verbal consent date	⦵	Feb 2020 ⌄
	Hospitalization	⦵	Feb 2020
			Hospitalized ⌄
☑	Smoking	⦵	Smoker ⌄

Get Data
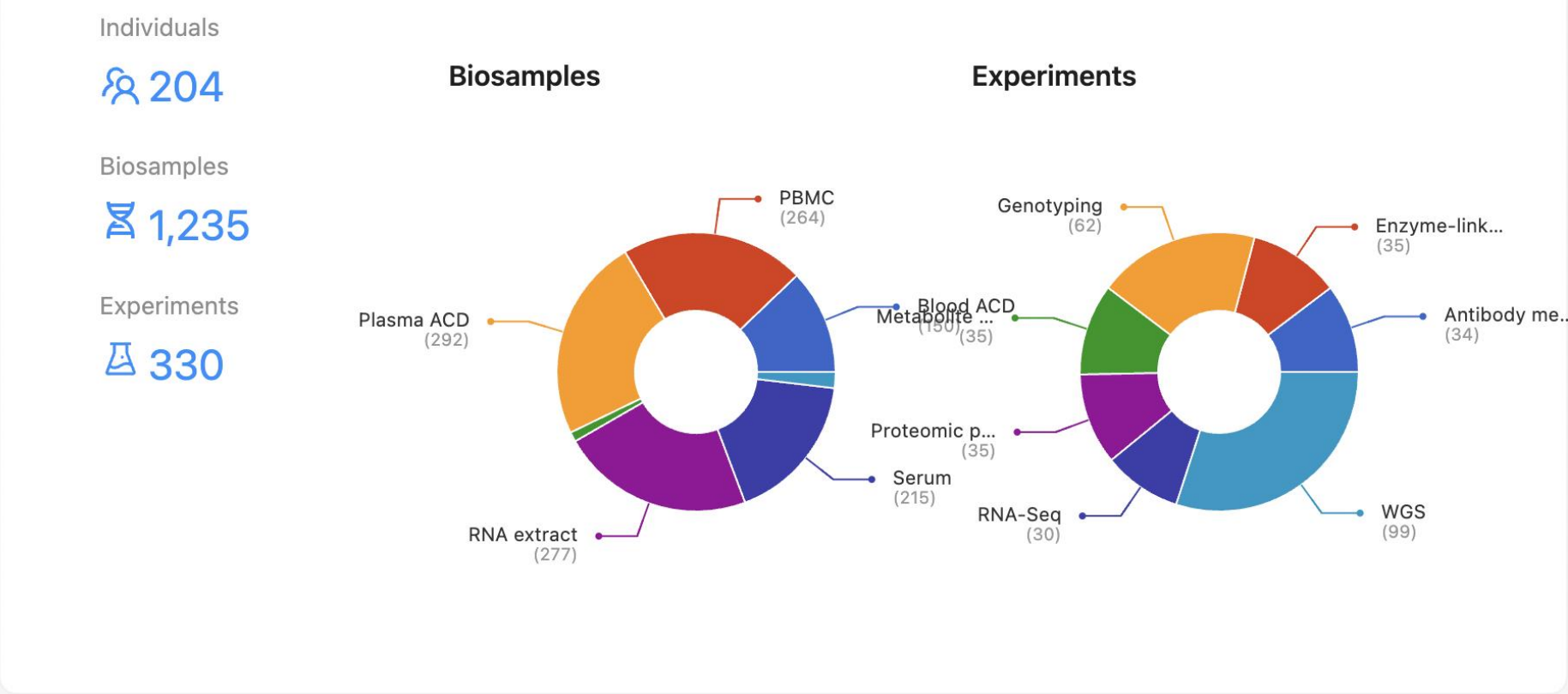
Individuals
🧑 128

Biosamples
⌛ 665

Experiments
⚗ 517

Biosamples

PBMC
(116)

Blood ACD
(132)

Metabolite ...
(66)

Plasma ACD
(190)

Serum
(63)

Proteomic p...
(68)

RNA extract
(146)

Experiments

Genotyping
(75)

Enzyme-link...
(66)

Antibody me...
(66)

WGS
(123)

RNA-Seq
(53)
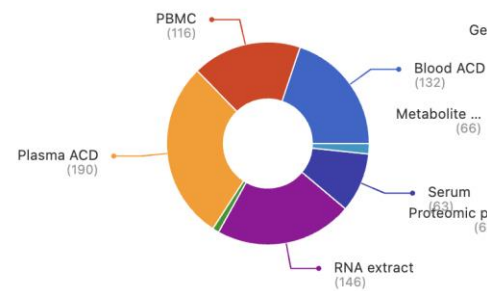
General

☐	Age	⊘	< 18
☐	Sex	⊘	FEMALE
☐	Verbal consent date	⊘	Feb 2020
☐	Hospitalization	⊘	Hospitalized
☐	Smoking	⊘	Smoker
☑	COVID status	⊘	Negative
☐	Vital status	⊘	Alive
☐	BMI	⊘	< 20
☑	ICU	⊘	Yes
☐	Participant category	⊘	Hospitalized adult
☐	Domicile	⊘	At home
☐	Vaccinal status	⊘	Yes
☐	Vaccine dose	⊘	1

Get Data

Individuals
👥 345

Biosamples
🧬 4,751

Experiments
⚗ 2,267

Biosamples

PBMC
(965)

Plasma ACD
(1297)

Blood ACD
(544)

Metabolite ...
(348)

Serum
(6...)
Proteomic p...
(338)

RNA extract
(1072)

Experiments

Genotyping
(210)

Enzyme-link...
(349)

Antibody me...
(336)

WGS
(288)

RNA-Seq
(398)

Proteomic p...
(338)
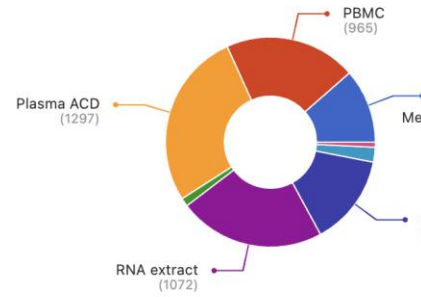
Treatments and outcomes

☐	Respiratory support	?	Yes ⌄
☐	Reinfection	?	Yes ⌄
☐	Persisting symptoms	?	Yes ⌄
☐	Systemic corticosteroid	?	Yes ⌄
☐	Antocoagulants	?	Yes ⌄
☐	Self-care post covid	?	Worse ⌄
☐	Acute kidney injury	?	Yes ⌄
☑	ARDS	?	Yes ⌄
☐	Deep vein thrombosis	?	Yes ⌄
☐	Stroke	?	Yes ⌄
☐	Pulmonary embolism	?	Yes ⌄

Get Data

FUTURE DEVELOPMENTS – BENTO 2.0

- Secure, password protected portal for users of BQC-19

- Deployment of analysis tools without need to download data

- Vizualisation tools at the patient level

- Larger quantity of search criteria available (presently limited to 2)

- No small cell rule

- Summary statistics by subgroups

- Allows to "cut" parts of data for subgroups for download

CONCLUSIONS - LESSONS LEARNED

- BQC-19 is one of many COVID biobanks containing wealth of data and biosamples that represent a rich tool for research: diverse populations and prolonged follow-up ideal to study waning and hybrid immunity, long COVID syndromes

- Large and complex datasets are hard to grasp: identification of sub-populations of interest to various researchers is important

- Bento offers a data visualization tool to promote **FAIR** principles (Findability, Accessibility, Interoperability, and Reusability)
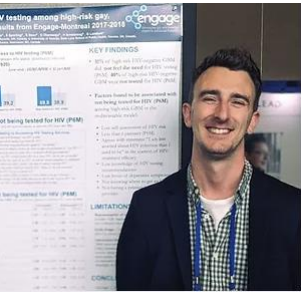
THANKS TO THE BQC19 CORE TEAM

Pascale Léon

Gestionnaire de la BQC19

Mylène Bertrand

Coordonnatrice de la BQC19

Marc Messier-Peet

Coordonnateur de la BQC19, éthique et convenance

David Bujold

Chef d'équipe de la gestion des données BQC19

Doris Ransy

Officière d'accès de la BQC19

Solomia Yanishevsky

Administratrice des données de la BQC19

Tania Abou Younes

Finance Manager

Finance Manager

QUESTIONS?