Pooling for SARS-CoV-2 control in care institutions

Jorge Julio Cabrera Alvargonzalez 1,2,3, Sonia Rey Cao 1,2, Sonia Pérez Castro 1,2,3, Lucía Martinez Lamas 1,2, Olaia Cores Calvo 1,2, Julio Torres Piñon 2, Jacobo Porteiro Fresco 4, Julio Garcia Comesaña 5 and Benito Regueiro Garcia 1,2,6*

Abstract

Background: Workers and residents in Care Homes are considered at special risk for the acquisition of SARS-CoV-2 infection, due to the infectivity and high mortality rate in the case of residents, compared to other containment areas. The role of presymptomatic people in transmission has been shown to be important and the early detection of these people is critical for the control of new outbreaks. Pooling strategies have proven to preserve SARS-CoV-2 testing resources.

The aims of the present study, based in our local experience, were (a) to describe SARS-CoV-2 prevalence in institutionalized people in Galicia (Spain) during the Coronavirus pandemic and (b) to evaluate the expected performance of a pooling strategy using RT-PCR for the next rounds of screening of institutionalized people.

Methods: A total of 25,386 Nasopharyngeal swab samples from the total of the residents and workers at Care Homes in Galicia (March to May 2020) were individually tested using RT-PCR. Prevalence and quantification cycle (Cq) value distribution of positives was calculated. Besides, 26 pools of 20 samples and 14 pools of 5 samples were tested using RT-PCR as well (1 positive/pool). Pooling proof of concept was performed in two populations with 1.7 and 2% prevalence.

Results: Distribution of SARS-CoV-2 infection at Care Homes was uneven (0–60%). As the virus circulation global rate was low in our area (3.32%), the number of people at risk of acquiring the infection continues to be very high. In this work, we have successfully demonstrated that pooling of different groups of samples at low prevalence clusters, can be done with a small average delay on Cq values (5 and 2.85 cycles for pools of 20 and 5 samples, respectively).

Conclusions: A new screening system with guaranteed protection is required for small clusters, previously covered with individual testing. Our proposal for Care Homes, once prevalence zero is achieved, would include successive rounds of testing using a pooling solution for transmission control preserving testing resources. Scale-up of this method may be of utility to confront larger clusters to avoid the viral circulation and keeping them operative.

Keywords: SARS-CoV-2, Pooling, Detection, Care homes, Low prevalence, Transmission control

* Correspondence: benito.jose.regueiro.garcia@sergas.es
1Microbiology and Infectology Research Group, Galicia Sur Health Research Institute (IIS Galicia Sur), SERGAS-UVIGO, Vigo, Spain
2Microbiology Department, Complexo Hospitalario Universitario de Vigo (CHUVI), Sergas, Vigo, Spain
Full list of author information is available at the end of the article
Background
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused more than 454,000 deaths since late 2019 [1]. Screening of Care Homes has been critical to limit the mortality rate in Galicia (Spain). Direct viral detection by real time RT-PCR was useful to identify people with potential SARS-CoV-2 transmission risk. Limited stocks and restrictions in test capacity prevented a higher number of RT-PCR tests per day.

Pooling strategies have proven to preserve SARS-CoV-2 testing resources and time with an increase in testing capability of the 69% for an incidence rate of SARS-CoV-2 infection of 10% or less [2–8], but it could be associated with a decrease in detection [9, 10]. Main limitations could be the preanalytical step, the sample viral load or the increase of the limit of detection of the individual sample [6].

The rationale in this study is to develop a new strategy based on initial individual identification of positive coronavirus cases in order to organize low prevalence clusters, followed by a serial pooling strategy testing of these clusters, in order to control areas free of virus circulation, allowing them to be fully operative.

Methods
Samples
Nasopharyngeal swab samples were obtained from residents and workers at Care Homes in Galicia (March to May 2020) and conserved in viral transport medium. The study protocol (2020/298) was approved by the Galician network of committees of research ethics.

Care homes screening
Samples were mixed 1:1 with cobas® omni lysis reagent (43% guanidine thiocyanate) for viral inactivation before individual testing. The Open Reading Frame (ORF) 1/b non-structural region of SARS-CoV-2 and the envelope E-gene of Sarbecovirus were detected with the cobas® SARS-CoV-2 test (Roche Diagnostics). For all RT-PCRs based on initial individual identification of positive coronavirus cases in order to organize low prevalence clusters, followed by a serial pooling strategy testing of these clusters, in order to control areas free of virus circulation, allowing them to be fully operative.

Results
SARS-CoV-2 screening of care homes
During the Coronavirus pandemic, SARS-CoV-2 prevalence was obtained by individually testing of 25,386 people from 306 Galician Care Homes: 16,477 residents, 8599 workers and 310 not specified. The mean age of workers and residents was 44.25 years (min 18, max 69) and 80.07 years (min 3, max 109), respectively. SARS-CoV-2 was detected in 852 people (3.32, 95% CI: 3.10–3.54%). The distribution of institutions by SARS-CoV-2 prevalence is shown in Fig. 1. A total of 282 institutions (21,861 people) had SARS-CoV-2 prevalence < 4%, including 263 institutions (19,091 people) with prevalence zero. Prevalence from 5 to 10% was observed in 2 institutions (389 people), from 10 to 20% in 11 institutions (1817 people) and from 20 to 60% in 11 institutions (1309 people).

Cq value distribution for positive samples was as follows: minimum 15.03/15.41, 1st quartile 21.86/22.55, median 26.41/27.54, 3rd quartile 31.36/33.60 and maximum 35.86/39.06 for ORF1b and E gene, respectively. Additional data of distribution and Cq values of positive samples were calculated.

Cq value distribution for positive samples was as follows: minimum 15.03/15.41, 1st quartile 21.86/22.55, median 26.41/27.54, 3rd quartile 31.36/33.60 and maximum 35.86/39.06 for ORF1b and E gene, respectively. Additional data of distribution and Cq values of positive samples detected during SARS-CoV-2 screening of Care Homes are available as additional files 1, 2 and 3.

Pooled positivity assessment
The selection of the optimal pool size should be made before the implementation of pooling testing. With non-overlapping pools, only positive pools will be retested. The reduction of the expected number of tests depends on the prevalence, the initial pool size and the number of stages for the pooling algorithm. In fact, it is generally accepted that 5% could be the prevalence threshold to
achieve a 50% reduction in the expected number of tests per individual. On the other hand, the sensitivity and specificity of the global process depends on the analytical characteristics of the test and on the number of times one sample is retested. Differences in the expected number of tests per individual, based on mathematical simulations, could help to choose the best set of pool sizes. According to other authors [11], for prevalence between 1 and 2%, sensitivity 95% and specificity 100%, the optimal pool size would be between 25 and 16 samples and the optimal sub pool size would be between 4 and 5 samples. In order to minimize the false negative factor for pooled testing recently defined [3] and to standardize the pooling method, pools of twenty samples (P20) and sub pools of five samples (SP5) were selected.

Test performance of twenty-six P20 and fourteen SP5 was studied. Each pool included one positive sample. A total of twenty-six positive samples were tested. Mean Cq values were 27.43 and 28.68 for ORF1/b and E gene, respectively. A boxplot of paired Cq values is shown in Fig. 2. All positive samples yielded a global positive result when tested in P20 or SP5. Sensitivity of E, RdRP and N gene was, respectively, 88.5% (23/26), 84.6% (22/26) and 96.1% (25/26) for P20. Sensitivity was 92.9% (13/14) for the three targets for SP5.

Mean delay in the Cq values (Cq pool– Cq positive sample) was 5.02 cycles for the P20 and 2.85 cycles for the SP5 (Table 1). An example of the amplification curves obtained for one particular sample is shown in additional files 4, 5 and 6. The N gene was not detected by Allplex™2019-nCoV assay in one specific sample independently of pooling or individual testing.

Proof of concept
Samples from Care Homes selected by prevalence were retrospectively tested in pools using the following algorithm: P20, SP5 when positive, individual analysis when positive. A first simulation was performed with 100 samples from 2% (95% CI: 0.24–7.04%, 2/100) prevalence Care Homes. Five P20 were tested. As 2 positive pools were obtained, 8 SP5 were processed. Two SP5 were positive, so 10 samples were tested individually. Two samples were positive. Number of tests was reduced 77% (0.23 tests per individual).

A second simulation included 60 samples from 1.7% (95% CI: 0.04–8.94%, 1/60) prevalence institutions. Three P20, 4 SP5 and 5 individual samples were tested. One sample was positive. Number of tests was reduced by 80% (0.20 tests per individual).

Discussion
A global SARS-CoV-2 seroprevalence of 5% in Spain [12] and a global viral prevalence around 3% at Care Homes reported in the present study, suggest that the number of people at risk of acquiring the infection continue to be very high. The role of transmission before symptoms has been shown to be important, presymptomatic / asymptomatic individuals may contribute to it [13, 14]. For these reasons their early detection seems critical to prevent further outbreaks. To control the spread of the virus, it is essential to detect as many infected individuals as possible, as quickly as possible to trace down and test possible contacts [15].

We performed the screening of 306 Care Homes (25, 386 determinations) in workers and residents using individual testing by RT-PCR. With a prevalence < 2% for
Table 1 Differences in quantification cycle (Cq) values between pooling and individual testing by Allplex™2019-nCoV assay

| Pool size | Parameter | Testing Condition | E gene Mean (range) | N gene Mean (range) | RdRP gene Mean (range) |
|-----------|-----------|-------------------|--------------------|---------------------|------------------------|
| P20 (n = 26) | Cq value | individual       | 24.63 (12.96, > 40) | 27.79 (15.4, > 40) | 26.23 (13.69, 33.45) |
|           |          | pool              | 29.67 (16.48, > 40) | 32.65 (19.9, > 40) | 31.39 (18.2, > 40)    |
|           |          | difference        | 5.04 (−1.06, 13.15) | 4.86 (−0.42, 11.52) | 5.16 (−1.39, 11.57)  |
| SP5 (n = 14) | Cq value | individual       | 21.38 (12.96, 28.48) | 25.92 (15.4, > 40) | 23.34 (13.69, 31.15) |
|           |          | pool              | 24.49 (14.5, > 40)  | 28.30 (18.23, > 40) | 26.42 (16.48, > 40)  |
|           |          | difference        | 3.11 (−2.26, 13.01) | 2.37 (−2.32, 6.66)  | 3.08 (−3.33, 10.22)  |

P20 Pool of 20 samples, SP5 sub pool of 5 samples, Cq Quantification cycle

Fig. 2 Boxplot of paired Cq values obtained for pooling and individual testing. Values are shown for each gene and pool size.
more than 85% people in Care Homes, pooling could achieve maximum usefulness. After reviewing the literature, and due to the absence of accumulated experience with this type of strategy for SARS-CoV-2, a more conservative pool size of 20 and sub pool size of 5 samples were chosen [2, 3, 6–8, 16].

Two tests authorized by the Food and Drug Administration were available at our laboratory. Both have shown suitable specificity and sensitivity for clinical diagnosis, but specific studies will be required for assessing their performance in pooling conditions. The choice of the Allplex™-2019-nCoV Assay was due to the flexibility and adaptability in the automation process useful for future interventions. Additionally, although it has been established a moderate mutation rate of SARS-CoV-2 [17–19], the possibility of detecting three targets could increase the possibilities of detection [18, 20].

As previous studies [9, 10, 16], our results using pools showed an increase of 3–5 cycles in the Cq value between pooled tests and individual positive samples. The pooling strategy was associated with a decreased sensitivity for individual targets (4–7%). Nevertheless, it has not carried out loss of global sensitivity in pools for samples included in this study. Samples of this study have been selected in order to represent those with Cq value within the first three quartiles observed in our population. A 100% global sensitivity was also achieved when testing Care Homes with prevalence around 2%, reducing until 80% the number of tests.

Proposed methodology
Here there is our proposal for introducing the pooling strategy for screening purposes in Care Institutions: When an institution with prevalence zero is characterized, successive rounds of pooling testing would be the option for transmission control. The maximum interval between rounds would be adjusted to avoid the loss of detection of infected people who could be in a phase of low viral load. The incubation period has been reported to be highly variable with an estimated average of 5–6 days [13, 21–24].

Limitations of this study were the limited number of samples included. Testing more negative samples would allow us to assess specificity and the risk of contamination along the processing. There is a likelihood of obtaining false negative results when a pooling strategy is introduced. Mainly, low quality samples cannot be discarded from the pools and the dilution could reduce the ARN concentration below the limit of detection. In this study we have focused on demonstrating that any pool containing until 20 individual samples from highly infectious people would be detected.

This work has shown the prevalence of SARS-CoV-2 in Spanish Care Homes during the Coronavirus pandemic. Prevalence differences shown between Institutions should address the interventions for viral transmission control. Few studies have assessed the performance of pooling for SARS-CoV-2 detection by rRT-PCR in real conditions, especially when aiming to keep areas free of virus circulation to be operative and functional.

Conclusions
Sample pooling could be a new testing strategy relevant for maintaining low level or no transmission among institutionalized people. Our proposal for Care Homes, once prevalence zero is achieved, would include successive rounds of testing using a pooling solution for transmission control preserving testing resources. Scale-up of this method may be of utility to confront larger clusters to avoid the viral circulation and keeping them operative. Further studies with self-sampling methods, modular systems and more specific pooling strategies will be necessary for the process improvement.

Supplementary information
Supplementary information accompanies this paper at https://doi.org/10.1186/s12879-020-05446-0.

Additional file 1. SARS-CoV-2 prevalence. Global prevalence is shown on the left. Stacked bar charts show Care Home prevalence obtained by individual testing for Care Homes with SARS-CoV-2 infections and without infection (prevalence zero).

Additional file 2. Distribution of SARS-CoV-2 RT-PCR Cq value. Summary of the distribution of Cq values of Care Homes with more than 5 positives. It is shown for each detected target. Samples were tested individually.

Additional file 3. Age Distribution. Distribution of age of Care Home residents and workers individually tested.

Additional file 4. E gene amplification curves. Example of amplification curves (E gene) obtained for the same sample processed individually and in pools of 5 (P5) and 20 (P20) samples. Obtained Cq values were 26.20, 29.31 and 30.82 for the individual sample, P5 and P20, respectively.

Additional file 5. N gene amplification curves. Example of amplification curves (N gene) obtained for the same sample processed individually and in pools of 5 (P5) and 20 (P20) samples. Obtained Cq values were 28.69, 31.91 and 33.26 for the individual sample, P5 and P20, respectively.

Additional file 6. RdRP gene amplification curves. Example of amplification curves (RdRP gene) obtained for the same sample processed individually and in pools of 5 (P5) and 20 (P20) samples. Obtained Cq values were 28.80, 31.49 and 31.79 for the individual sample, P5 and P20, respectively.

Abbreviations
SARS-CoV-2: Severe acute respiratory syndrome coronavirus-2; RT-PCR: Reverse transcription polymerase chain reaction; Cq: quantification cycle; ORF: Open reading frame; P20: Pool of twenty samples; SP5: Subpool of five samples.

Acknowledgements
To the Servizo Galego de Saúde, Consellería de Sanidade and Xunta de Galicia for supporting this study. To Dr. Vázquez Almuñé for his suggestions. To all the professionals who collaborated in information management, sampling, supplies, contact tracing and organization at all levels. Especially to the technician staff of the Microbiology Department of the CHUVI. To Laura Regueiro for her assistance with the manuscript.
Authors’ contributions
JG-C and BJR have coordinated with Consellería de Sanidade the screening in care institutions. BJR, JG-C, JP, JT, JJC, SR, SP participated in the study design. BJR, JJC and SP wrote the first draft. JJC, SR, SP, LM-L, OC-C have processed samples, collected data, participated in the analyses and reviewed the draft. All authors agreed on the final version of the manuscript.

Funding
No applicable, no external funding.

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. All data and materials were obtained under Servizo Galego de Saúde, Consellería de Sanidade de Xunta de Galicia (institutions belonging to our National Health Public System) and are under their regulations.

Ethics approval and consent to participate
Informed consent was not obtained since the samples were remnants of the clinical diagnosis (no additional samples were obtained for this study) and personal information was pseudonymized by people unrelated to the study investigators. This study was approved by Galician network of committees of research ethics (protocol number 2020-298).

Consent for publication
No applicable.

Competing interests
The authors: No reported conflicts of interest.

Author details
1Microbiology and Infectology Research Group, Galicia Sur Health Research Institute (IIS Galicia Sur), SERGAS-LUVIGO, Vigo, Spain. 2Microbiology Department, Complexo Hospitalario Universitario de Vigo (CHUVI), Sargas, Vigo, Spain. 3Universidade de Vigo, Vigo, Spain. 4CINTECX, Universidade de Vigo, GTE, Vigo, Spain. 5Management Department, Complexo Hospitalario Universitario de Vigo (CHUVI), Sargas, Vigo, Spain. 6Microbiology and Parasitology Department, Medicine and Odontology, Universidade de Santiago, Santiago de Compostela, Spain.

Received: 14 July 2020 Accepted: 23 September 2020

Publisher online: 12 October 2020

References
1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis. 2020; [cited 12 de mayo de 2020]; Available en: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7190186/.
2. Abdalhamid B, Bilder CR, McCutchen EL, Hinrichs SH, Koepsell SA, Iwen PC. Genomic Analysis of SARS-CoV-2 in Ten Patients from a Mid-sized City outside of Hubei, China. Infect Dis (except HIV/AIDS). 2020; abr [citado 21 de mayo de 2020]. Disponible en: http://medrxiv.org/lookup/doi/10.1101/2020.04.16.20058560.
3. Wang C, Liu Z, Chen Z, Huang X, Xu M, He T, et al. The establishment of reference sequence for SARS-CoV-2 and variation analysis. J Med Virol. 2020.
4. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The incubation period for coronavirus disease (COVID-19) based on all reported confirmed cases outside of Hubei, China. Infect Dis (except HIV/AIDS). 2020; abr [citado 21 de mayo de 2020]. Disponible en: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30246-5.
5. Hogan CA, Sahoo MK, Pinsky BA. Sample Pooling as a strategy to detect viral shedding and transmissibility of COVID-19. Nat Med mayo de. 2020; 26(5):672–5.
6. Yelin I, Aharony N, Shaer Tamar E, Argoetti A, Messer E, Berenbaum D, et al. Evaluation of Group Testing for SARS-CoV-2 Rna. medRxiv. 2020:2020.03.27.20043968.
7. Ganyani T, Kremer C, Chen D, Torneri A, Faes C, Wallinga J, et al. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. Eurosurveillance. 2020;25(17):2000257.
8. Ben-Ami R, Klochendler A, Seidel M, Sido T, Gurel-Gurevich O, Yassour M, et al. Pooling of samples for testing for SARS-CoV-2 in asymptomatic people. Lancet Infect Dis. 2020; [cited 14 de mayo de 2020](06). Disponible en: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30246-5.
9. Riou J, Althaus CL. Pattern of early human to human transmission of Wuhan novel coronavirus (2019-nCoV) December 2019 to January 2020. Euro Surveill. 2020;25(4):200058.
10. Tornatore E, Colaianni M, Yamauchi M, et al. A note on the effective reproduction number of COVID-19. medRxiv. 2020:2020.04.28.20083337.
11. ChrisBilder.com [Internet]. [cited 14 de mayo de 2020]. Disponible en: http://medrxiv.org/lookup/doi/10.1101/2020.04.16.20058560.
12. Delling G, Gnerlich H, Bentien C, Kappes J, Mata R, Kuchenbauer N, et al. Reported Confirmed Cases: Estimation and Application. Ann Intern Med. 2020;25(5).
13. Ben-Ami R, Pasternak MA, Iwen PC. A note on the effective reproduction number of COVID-19. medRxiv. 2020:2020.04.28.20083337.
14. Chen J, Hilt E, Wu H, Jiang Z, Zhang Q, Wang J, et al. Epidemiological and Genomic Analysis of SARS-CoV-2 in Ten Patients from a Mid-sized City outside of Hubei, China. Infect Dis (except HIV/AIDS). 2020; abr [citado 21 de mayo de 2020]. Disponible en: http://medrxiv.org/lookup/doi/10.1101/2020.04.16.20058560.
15. Wang C, Liu Z, Chen Z, Huang X, Xu M, He T, et al. The establishment of reference sequence for SARS-CoV-2 and variation analysis. J Med Virol. 2020.
16. Li X, Zai j, Zhao Q, Nie Q, Li Y, Foley BT, et al. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. J Med Virol. 2020.
17. Chen J, Hilt E, Wu H, Jiang Z, Zhang Q, Wang J, et al. Epidemiological and Genomic Analysis of SARS-CoV-2 in Ten Patients from a Mid-sized City outside of Hubei, China. Infect Dis (except HIV/AIDS). 2020; abr [citado 21 de mayo de 2020]. Disponible en: http://medrxiv.org/lookup/doi/10.1101/2020.04.16.20058560.
18. Wang C, Liu Z, Chen Z, Huang X, Xu M, He T, et al. The establishment of reference sequence for SARS-CoV-2 and variation analysis. J Med Virol. 2020.
19. Attuor S, Bontems S, Gobbel R, Frank R, Bexius D, Arten S, et al. Failure of the cobas® SARS-CoV-2 (Roche) E-gene assay is associated with a C-to-T transition at position 26340 of the SARS-CoV-2 genome. medRxiv. 2020:2020.04.28.20083337.
20. Ben-Ami R, Pasternak MA, Iwen PC. A note on the effective reproduction number of COVID-19. medRxiv. 2020:2020.04.28.20083337.
21. Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung S-M, et al. Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data. J Clin Med. 2020;2(2).
22. Lauer SA, Grantz KH, Bi Q, Jaimes JT, Zheng Q, Meredith HR, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. Ann Intern Med. 2020; [cited 21 de mayo de 2020]. Disponible en: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7081172/.
23. Lauer SA, Grantz KH, Bi Q, Jaimes JT, Zheng Q, Meredith HR, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. Ann Intern Med. 2020; [cited 21 de mayo de 2020]. Disponible en: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7081172/
24. Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 viral Load in Upper Respiratory Specimens of Infected Patients. N Engl J Med. 2020;382(12):1177-9.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.