

Group Testing with Homophily to Curb Epidemics with Asymptomatic Carriers

Louis-Marie Harpedanne de Belleville

Paris School of Economics, Université Paris 1 Panthéon-Sorbonne and Banque de France

October 31, 2020

Summary

Contagion happens through heterogeneous interpersonal relations (homophily) which induce contamination clusters. Group testing is increasingly recognized as necessary to fight the asymptomatic transmission of the COVID-19. Still, it is plagued by false negatives. Homophily can be taken into account to design test pools that encompass potential contamination clusters. I show that this makes it possible to overcome the usual information-theoretic limits of group testing, which are based on an implicit homogeneity assumption. Even more interestingly, a multiple-step testing strategy combining this approach with advanced complementary exams for all individuals in pools identified as positive identifies asymptomatic carriers who would be missed even by costly exhaustive individual tests. Recent advances in group testing have brought large gains in efficiency, but within the bounds of the above cited information-theoretic limits, and without tackling the false negatives issue which is crucial for COVID-19. Homophily has been considered in the contagion literature already, but not in order to improve group testing.

1 I. INTRODUCTION

2 Massive and timely identification of asymptomatic disease carriers is crucial if human-to-
3 human asymptomatic transmission happens, which is documented for COVID19^{1 2 3 4 5}. Li, Pei
4 *et al.* (2020)⁶ find that although the transmission rate of undocumented carriers is only 55% that
5 of documented carriers, the former are responsible for 80% of contaminations.

6 Massive and repeated RT-PCR testing is possible only through group testing (testing a pool of
7 swabs of many individuals). Group testing is used to fight COVID19 in China, India, Germany,
8 the United States⁷ and Rwanda⁸. For the literature on group testing and COVID19, see ^{7 8 9 10 11}
9 ^{12 13 and 14} With a .1 percent prevalence, the two-step adaptive design proposed by Dorfman
10 (1943)¹⁵ decreases 17-fold the number of tests required to identify asymptomatic COVID-19
11 carriers (0.06 test per person) while the strategy suggested by Mutesa *et al.* (2020)⁸ decrease it
12 55 times (0.018 test per person). Still, these testing strategies are plagued by false negatives
13 (see Section IV).

14 Contagion happens through heterogeneous interpersonal relations (homophily) which can be
15 identified *ex ante* (Section II) to design test pools that encompass potential contamination
16 clusters. Thus, it is possible to overcome information-theoretic limits, which rely on an implicit
17 homogeneity assumption, and to make tests more efficient (Section III). Combining this
18 approach with individual complementary exams for the positive groups identifies carriers who
19 would be missed even by costly exhaustive individual tests on nasopharyngeal swabs
20 (Section IV).

21 II. HOMOPHILY IS PREVALENT AND CAN BE IDENTIFIED EX ANTE

22 The social sciences and epidemiology literatures show that heterogeneous interpersonal
23 interactions, which induce small potential contamination clusters, are recurrent and can be
24 identified *ex ante*. Thus, it is possible to design test pools encompassing potential clusters.

25 Homophily, defined in 1954 by Lazarsfeld and Merton¹⁶, “refers to the fact that people are more
26 prone to maintain relationships with people who are similar to themselves”¹⁷. Homophily is
27 prevalent in many social networks¹⁸ and affects contagion¹⁹. Moulton (1986, 1990)^{20,21},
28 introduced a related econometric concept, clustering, which refers to the nondeterministic
29 correlation of outcomes between individuals that are somewhat related: failing to take it into
30 account induce significant errors when estimating standard errors. Clustering has been
31 popularized by Bertrand *et al.* (2004)²² and extended to multiple non-nested dimensions²³.
32 Correcting for potential clustering is a condition *sine qua non* for publication in applied
33 economics.

34 The epidemiologic literature confirms the importance of clusters. Han and Yang (2020)²⁴ cite a
35 Chinese-written article asserting that “In some cities, cases involving cluster transmission
36 accounted for 50% to 80% of all confirmed cases of COVID-19.” According to a meta-analysis
37 of 20 studies, households display high secondary attack rate (15.4% on average)²⁵, and out of
38 36 children infected in a Chinese city, 32 (89%) had transmission by close contact with family
39 members²⁶. High SARs also occurs in a chalet (73.3%), at a choir (53.3%), at a religious event,
40 or for travels and eating with an index case. Other cases of clusters with very high absolute
41 attack rate include “a nursing home in Kings County, Washington (64%) [...] a church in
42 Arkansas (38%), a homeless shelter in Boston (36%), a fitness dance class (26.3%) and the
43 Diamond Princess cruise ship in Japan (18.8%)”²⁵. Park *et al.* (2020)²⁷ analyze an outbreak in
44 a building: 94 of 97 cases worked on the same floor, and 79 in the same open-space (attack rate
45 of 52%). Many clusters have been observed in slaughterhouses.

46 *Ex ante* general contamination patterns in small clusters can be identified. Longer and more
47 intense exposure increases the risk of infection^{25 28 29 30}; so do indoor environments with
48 sustained close contact and conversations³¹. In a call center, cases are concentrated in large
49 open spaces but only one case in small offices²⁷. This is consistent with Harpedanne (2020)³²

50 who shows a convex theoretical contamination effect of the number of users. Using these
51 general patterns and theoretical results, it is possible to identify *ex ante* potential clusters, and
52 to design pools that encompass these clusters. Sections III and IV quantify the gains from this
53 strategy.

54 III. INFORMATION THEORETIC LIMITS CAN BE OVERCOME WHEN DESIGNING 55 TEST POOLS

56 Here, I show that homophily contains information that can make group testing more efficient
57 theoretically. For that purpose, I show that strong homophily used to design the testing pools
58 makes it possible to overcome the most recent and tight information-theoretic lower bounds on
59 the efficiency of group testing, identified by Chan *et al.* (2011)³³ who claim to be the first in
60 the literature to define limits in terms of actual numbers and not only rate or capacity, and
61 Baldassini *et al.* (2013)³⁴, who follow the same path and provide a new and tighter lower bound.

62 I focus here on noiseless tests: a negative test outcome is guaranteed when all items in the
63 testing pool are nondefective, and a positive outcome when a least one item in the pool is
64 defective³⁵. Otherwise, the test is noisy.

65 Noisy tests are often examined under the assumptions of constant³³ or worst-case³⁶ noise.
66 However, the literature points instead to a risk of false negatives increasing with dilution¹⁴
67 (analyzed in a previous draft), and to patient-specific idiosyncratic noise^{37 38 39}. General form
68 noise models (the symmetric error model³³ or the additive model⁴⁰) are irrelevant to analyze
69 idiosyncratic noises. Also, focusing here on noiseless tests shows that the benefits of taking
70 homophily into account in group testing are not limited to noise-related issues, and makes the
71 comparison with the information-theoretic limit of Baldassini *et al.* (2013)³⁴ easier.
72 Furthermore, Baldassini *et al.* (2013, Section III)³⁴ analyze a noiseless test in a population of
73 size N , with K defectives (K is known for simplicity). They show that if the number of tests is
74 limited to T , the probability of correct identification of the set of defectives is:

75 (1)
$$P(suc) \leq \frac{2^T}{C_N^K}$$

76 where $C_N^K = \frac{N!}{K!(N-K)!}$.

77 Let build a simple counterexample. N is 64 and K is 8. Using 6 tests only ($T=6$), one can cut the
78 population in 8 groups of 8 people each and determine which group contains carriers *if only one*
79 *group contains carriers* (think of the 64 population as a 4x4x4 cube and cut it in half in each
80 dimension, that is implement 6 tests over 32 people each). According to (1):

81 (2)
$$P(suc) \leq \frac{2^6}{C_{64}^8} \approx 1.45 \cdot 10^{-8}$$

82 With homophily, there exists high potential for within-group contamination and low potential
83 for between-groups contamination. With probability $(1 - \varepsilon_1)$, only one individual has imported
84 the disease in the 64 population (a decent assumption with a low general prevalence); ε_2 is the
85 probability that intergroup contamination has happened. Then with probability $(1 - \varepsilon_1)(1 - \varepsilon_2)$,
86 all 8 carriers are in the same group. Let fix $\varepsilon_1=0.2$ and $\varepsilon_2=.5$. Then:

87 (3)
$$P(suc) \geq 0.4$$

88 which contradicts (2): homophily provides information that makes it possible to overcome
89 information-theoretic limits based on an implicit homogeneity assumption.

90 IV. FALSE NEGATIVES CAN BE IDENTIFIED USING A TWO-STEP STRATEGY

91 The swabs of many disease carriers fail to contain viral loading, inducing patient-specific
92 idiosyncratic false negatives^{37 38 39}. A strategy based on group testing with homophily can solve
93 this issue and identify more asymptomatic carriers than exhaustive individual testing.

94 Methods to identify SARS-CoV-2 carriers include clinical diagnosis (not for asymptomatic
95 carriers), chest radiograph and CT-scan (not very available), fibrobronchoscope brush biopsy
96 and RT-PCR on bronchoalveolar lavage fluid (requires specific equipment and skilled
97 operators), sputum (produced in only 28 % of the COVID cases⁴¹), feces swabs, nasal and throat

98 swabs... Only nasal and throat swabs can be used easily for large scale asymptomatic testing.
99 The former induce a limited rate of false negatives if implemented less than one week after the
100 onset of the disease for symptomatic patients (table 1). Thus, a strategy based on nasal swabs
101 should be repeated weekly to minimize the risk of false negatives.

102 **Table 1 : rate of false negatives on RT PCR tests**

RT PCR on :	Wang, Tan, Wang et al. (2020)	Yang, Yang, Shen <i>et al.</i> (2020) mild cases			Wang, Xu, Gao et al (2020)
	inpatients	< 7 d.a.o.	8~14 d.a.o.	> 15 d.a.o	
Throat	90.7 %	38.7 %	70.4 %	88.9 %	68 %
Nasal	67.1 %	27.9 %	46.4 %	45.5 %	37 %
<i>Throat & nasal</i>	<i>64.6 %</i>	<i>NA</i>	<i>NA</i>	<i>NA</i>	<i>NA</i>
Sputum	NA	17.8 %	25.6 %	57.1 %	25 %
BALF	NA	100 %	100 %	100 %	7 %
Fibrobronchoscope brush biopsy	NA	NA	NA	NA	54 %
Feces	NA	NA	NA	NA	71 %
Blood	NA	NA	NA	NA	99 %
Urine	NA	NA	NA	NA	100 %

103 Still, even nasal swabs, widely used for identification of SARS-CoV-2, display a significant
104 rate of false negatives. Group testing with homophily is very beneficial here. If a carrier is “false
105 negative” with probability α , then a pool with two carriers will be negative with probability α^2
106 only ($\alpha < 1$), and with probability α^n if there are n carriers: concentrating carriers in a pool
107 decrease the risk of wrongly classifying this pool as negative.

108 In a second step, implementing individual RT-PCR on nasopharyngeal swab would reintroduce
109 a false negative problem. Rather, advanced complementary exams are implemented on all
110 individuals belonging to positive pools: chest radiograph or CT scan, fibrobronchoscope brush
111 biopsy, and RT-PCR on additional types of swabs. With independence, the probability that all
112 tests provide a false negative would be the product of the proportion of false negatives for each
113 type of test. Empirical evidence on this issue is scarce: The correlation between nasal and throat
114 swabs is low (Kappa=0.308) and Computed Tomography scan was always able to detect
115 ground-glass opacities for cases without viral shedding in the swabs examined (3 cases)³⁹.
116 Similar results were obtained for PiO₂/FiO₂ and Murray score, implemented on two cases

117 without viral shedding, always pointed to lung injury. Whether these results apply to
118 asymptomatic carriers is an open question. More advanced work is needed on the correlation
119 between these tests and exams for asymptomatic carriers, but a series of tests including nasal,
120 throat and feces swabs, sputum swabs when available, lower respiratory swabs, PiO₂/FiO₂ tests
121 and Murray score, may be able to detect most individual carriers if implemented frequently; CT
122 scan may prove useful to identify more severe case.

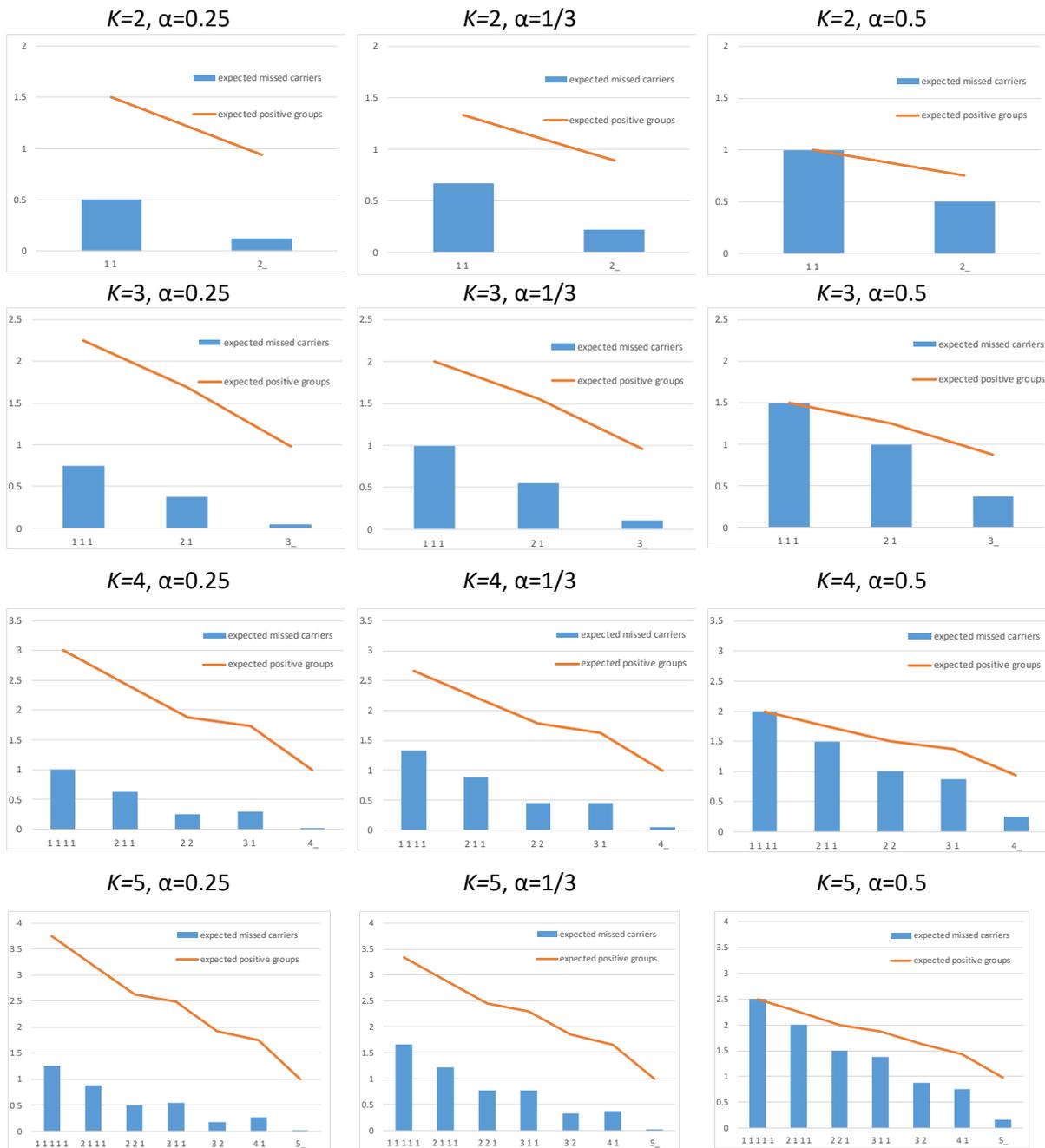
123 Graph 1 analyze the quantitative gains from homophily in this two-step strategy, for 2 to 5
124 defectives. This covers a large range of situations: two defectives out of two thousand people
125 correspond to a rate of 0.1%, while five out of 50 correspond to 10%. The size (and number) of
126 pools do not affect the graphs, which therefore cover a large range of pool size.

127 From left to right, each graph displays increasing concentration (denoted by $<_C$) of the
128 defectives in testing pools. $<_C$ is transitive but is not a total order, and using brackets \approx for non-
129 ordered configurations, we get, for 5 defectives:

130 (4) $(1\ 1\ 1\ 1\ 1) <_C (2\ 1\ 1\ 1) <_C (2\ 2\ 1) \approx (3\ 1\ 1) <_C (3\ 2) \approx (4\ 1) <_C (5)$.

131 Bars describe the expected number of missed carriers due to false negatives. When carriers are
132 concentrated in a few pools or a single pool, the expected number of missed carriers decreases.
133 The gains from homophily can be summarized by comparing the expected number of missed
134 carriers when all carriers are in different pools and when they are in the same pool. For $\alpha=1/3$
135 (which is the central case^{37 39}), the reduction reaches 67% for two carriers (0.67 expected missed
136 carriers if the carriers are in two different groups against 0.222 if both are in the same pool),
137 89% for three carriers, 96% for four carriers and 98.7% for five carriers. The absolute and
138 relative gains are higher for $\alpha=0.5$ (respectively 75%, 94%, 98.4% and 99.6% for two, three,
139 four and five carriers) and lower for $\alpha=0.25$ (respectively 50%, 75%, 87.5% and 93.75%). Even
140 less successful concentration can still bring huge gains. For instance, with five carriers and
141 $\alpha=1/3$, getting three carriers in a pool and two in another brings a reduction of 80%.

Graph 1: benefits of group testing and homophily to confront individual false negative



142 Note that even without homophily, that is if carriers are *i.i.d.* in the different pools, two carriers
 143 may be in the same pool by chance, which reduces the expected number of “missed” carriers.
 144 Group testing even without homophily can help identify and isolate carriers who would be
 145 missed by exhaustive individual testing. To the best of my knowledge, this simple and striking
 146 fact has been overlooked in the scientific and policy debates on group testing to fight epidemics.

147 Advanced complementary exams are costly. Homophily is crucial here. When the concentration
148 of carriers increases, the expected number of contaminated groups, and therefore the expected
149 number of positive groups (groups that are *identified* as contaminated by the first step) are
150 reduced stringently. The gains of homophily are lower than those observed for the expected
151 number of missed carriers, because homophily has two opposite effects here: concentration
152 decreases the number of potential positive groups, but improves the identification of these
153 groups. Still, these gains are significant. For $\alpha=1/3$, the reduction in the number of groups on
154 which to implement complementary exams is 33.3% for two carriers, 51.8% for three carriers,
155 63% for four carriers and 70% for five carriers. The absolute gains are especially high for low
156 values of α , since for low α , most contaminated groups are identified correctly and homophily
157 reduces mainly the number of really contaminated groups.

158 V. CONCLUSION

159 Homophily is prevalent in interpersonal interactions. Taking it into account makes group testing
160 more efficient and help reduce false negatives issues. In case of a lockdown, testing together
161 households, which are all potential clusters with high attack rate, is efficient; in more normal
162 times, a multi-dimensional testing strategy with non-nested pools (households on the one hand,
163 other potential cluster such as firms on the other hand) is likely to be beneficial. These results
164 open a new avenue for research to fine tune the present analysis and combine it with other
165 research in order to better fight the COVID19 epidemic.

166

167 The author declares no competing interest.

168

169 Acknowledgements

170 I thank Marc Fleurbaey, Matthew Jackson, Xavier d'Haultfeuille and seminar participants at
171 Paris School of Economics for useful comments. All remaining errors are mine.

172 REFERENCES

- ¹ Bai, Y., Yao, L., Wei, T., Tian, F., Jin, D.-Y., Chen, L., and Wang, M. (2020), “Presumed asymptomatic carrier transmission of COVID-19,” *JAMA*. (Epub ahead of print)
- ² Rothe, C., Schunk, M., Sothmann, P., Bretzel, G., Froeschl, G., Wallrauch, C., Zimmer, T., Thiel, V., Janke, C., Guggemos, W., Seilmaier, M., Drosten, C., Vollmar, P., Zwirgmaier, K., Zange, S., Wölfel, R., Hoelscher, M., (2020, March 5), “Transmission of 2019-nCoV infection from an asymptomatic contact in Germany,” *New England Journal of Medicine* (Epub ahead of print)
- ³ Zou L., Ruan, F., Huang, M., Liang, L., Huang, H., Hong, Z., Yu, J., Kang, M., Song, Y., Xia, J., Guo, Q., Song, T., He, J., Yen, H. L., Peiris, M., and Wu, J. (2020, March 19), “SARS-CoV-2 viral load in upper respiratory specimens of infected patients,” *New England Journal of Medicine*.
- ⁴ Santarpia, J. L., Rivera, D. N., Herrera, V., Morwitzer, M. J., Creager, H., Santarpia, G. W., Crown, K. K., Brett-Major, D. M., Schnaubelt, E., Broadhurst, M. J., Lawler, J. V., St. Reid, P., and Lowe, J. J. (2020, March 26), “Transmission potential of SARS-CoV-2 in viral shedding observed at the University of Nebraska Medical Center,” *mimeo*.
- ⁵ WONG, Justin, Anita B Z ABDUL AZIZ, Liling CHAW , Abdirahman MAHAMUD, Matthew M. GRIFFITH, Ying-Ru LO, Lin NAING (2020) “High proportion of asymptomatic and presymptomatic COVID-19 infections in travelers and returning residents to Brunei”, *Oxford University Press – Public Health Emergency Collection*.
- ⁶ Li, Ruiyun, Sen Pei , Bin Chen, Yimeng Song, Toa zang, Wan Yang and Jeffrey Shaman (2020), “Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2)”, *Science*, 368:489-493. doi:[10.1126/science.abb3221](https://doi.org/10.1126/science.abb3221)
- ⁷ Mallapaty, Smriti (2020), “The mathematical strategy that could transform coronavirus testing”, *Nature* (10 July 2020). <https://www.nature.com/articles/d41586-020-02053-6>
- ⁸ Mutesa, Leon, Pacifique Ndishimye, Yvan Butera, Jacob Souopgui, Annette Uwinezwa, Robert Rutayisire, Ella Larissa Ndoricimpaye, Emile Musoni, Nadine Rujeni, Thierry Nyatanyi, Edouard Ntagwabira, Muhammed Semakula, Clarisse Musanabaganwa, Daniel Nyamwasa, Maurice Ndashimye, Eva Ujeneza, Ivan Emile Mwikarago, Claude Mambo Muvunyi, Jean Baptiste Mazarati, Sabin Nsanzimana, Neil Turok & Wilfred Ndifon (2020) “A pooled testing strategy for identifying SARS-CoV-2 at low prevalence”, *Nature*, <https://www.nature.com/articles/s41586-020-2885-5>
- ⁹ Gollier, Christian and Olivier Gossner (2020), “Group Testing against COVID-19”, *CREST DT* 2020-04.
- ¹⁰ Conger, K. (2020, April). Testing pooled samples for COVID-19 helps Stanford researchers track early viral spread in Bay Area. Stanford Medicine News Center.
- ¹¹ Lakdawalla, Darius, Emmett Keeler, Dana Goldman, Erin Trish (2020), “Getting Americans Back to Work (and School) with Pooled Testing”, USC Schaeffer Center White Paper.
- ¹² Eberhardt, Jens Niklas, Nikolas Peter Breuckmann and Christiane Sigrid Eberhardt (2020), “Multi-Stage Group Testing Improves Efficiency of Large-Scale COVID-19 Screening”, *Journal of Clinical Virology* doi: [10.1016/j.jcv.2020.104382](https://doi.org/10.1016/j.jcv.2020.104382)
- ¹³ Lohse, Stefan ,Thorsten Pfuhl, Barbara Berkó-Göttel, Jürgen Rissland, Tobias Geißler, Barbara Gärtner,.., Sören L Becker, Sophie Schneitler, Sigrun Smola (2020, April 28), “Pooling of Samples for testing for SARS-CoV-2 in asymptomatic people,” *Lancet Infectious Diseases*. [https://doi.org/10.1016/S1473-3099\(20\)30362-5](https://doi.org/10.1016/S1473-3099(20)30362-5).
- ¹⁴ Yelin, Idan, Noga Aharoni, Einat Shaer Tamar, Amir Argoetti, Esther Messer, Dina Berenbaum, Einat Shafran, Areen Kuzli, Nagam Gandali, Tamar Hashimshony, Yael Mandel-Gutfreund, Michael Halberthal, Yuval Geffen, Moran Szwarcwort-Cohen and Roy Kishony

(2020), “Evaluation of COVID-19 RT-qPCR test in multi-sample pools”, *Clinical Infectious Diseases*. doi: 10.1093/cid/ciaa531.

¹⁵ Dorfman, R. (1943), “The detection of defective members of large population”, *The Annals of Mathematical Statistics*, 436-440.

¹⁶ Lazarsfeld, Paul and Robert K. Merton (1954), “Friendship as a Social Process: A Substantive and Methodological Analysis”, pp. 18–66 in *Freedom and Control in Modern Society*, edited by M. Berger, T. Abel, and C. H. Page. New York: [Van Nostrand](#).

¹⁷ Jackson, Matthew O. (2008), “Social and Economic Networks”, *Princeton University Press*, Princeton, New Jersey, 504 pp.

¹⁸ McPherson, Miller; Lynn Smith-Lovin and James M. Cook (2003). "Birds of a Feather: Homophily in Social Networks", *Annual Review of Sociology*, Vol. 27 (1), pp. 415–444.

¹⁹ Jackson, M. O., and Lopez-Pintado, D. (2013), “Diffusion and contagion in networks with heterogeneous agents and homophily,” *Network Science* 1(1):49-67.

²⁰ Moulton, Brent R. (1986), “Random Group Effects and the Precision of Regression Estimates”, *Journal of Econometrics*, 32: 385-397.

²¹ Moulton, Brent R. (1990), “An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units”, *Review of Economics and Statistics* 72(3): 334-338.

²² Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan (2004), “How Much Should We Trust Difference-in-Difference Estimates?”, *The Quarterly Journal of Economics*, Vol. 119 (1), pp. 249–275

²³ Cameron, A. Colin., J. G. Gelbach and Douglas L. Miller (2011), “Robust Inference with Multi-Way Clustering”, *Journal of Business and Economic Statistics* 29(2): 238-249.

²⁴ Han, Yu and Hailan Yang (2020), “The transmission and diagnosis of 2019 novel coronavirus infection disease (COVID-19): A Chinese perspective”, *Journal of Medical Virology* 92, 639-644.

²⁵ Koh, Wee Chian L, in Naing, Muhammad Ali Rosledzana, Mohammad Fathi Alikhan, Liling Chaw, Matthew Griffith, Roberta Pastore and Justin Wong (2020), “What do we know about SARS-CoV-2 transmission? A systematic review and meta-analysis of the secondary attack rate, serial interval, and asymptomatic infection”, *medRxiv*.

²⁶ Qiu H, J. Wu, L. Hong, Y. Luo, Q. Song and D. Chen (2020) « Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: an observational cohort study. *Lancet Infectious Disease* doi:[10.1016/S1473-3099\(20\)30198-5](https://doi.org/10.1016/S1473-3099(20)30198-5)

²⁷ Park, Shin Young , Young-Man Kim, Seonju Yi, Sangeun Lee, Baeg-Ju Na, Chang Bo Kim, Jung-il Kim, Hea Sook Kim, Young Bok Kim, Yoojin Park, In Sil Huh, Hye Kyung Kim, Hyung Jun Yoon, Hanaram Jang, Kyungnam Kim, Yeonhwa Chang, Inhye Kim, Hyeyoung Lee, Jin Gwack, Seong Sun Kim, Miyoung Kim, Sanghui Kweon, Young June Choe, Ok Park, Young Joon Park, Eun Kyeong Jeong (2020), “Coronavirus Disease Outbreak in Call Center, South Korea”, *Emerging Infectious Diseases* 26 (8).

²⁸ Li W, Zhang B, Lu J, Liu S, Chang Z, Cao P, *et al.* (2020) “The characteristics of household transmission of COVID-19”. *Clinical Infectious Disease*; ciaa450.

²⁹ Madewell, Zachary J., Yang Yang, Ira M. Longini Jr, M. Elizabeth Halloran, Natalie E. Dean “Household transmission of SARS-CoV-2: a systematic review and meta-analysis of secondary attack rate”, *medRxiv*.

³⁰ Liu T, Liang W, Zhong H, *et al.* (2020) “Risk factors associated with COVID-19 infection: a retrospective cohort study based on contacts tracing”. *Emerging Microbes & Infection*: 1-31.

³¹ Nishiura H, Oshitani H, Kobayashi T, *et al.* (2020) “Closed environments facilitate secondary transmission of coronavirus disease 2019 (COVID-19)”. *medRxiv*.

³² Harpedanne de Belleville, Louis-Marie (2020), *Act Now or Forever Hold Your Peace: Slowing contagion with Unknown Spreaders, Limited Cleaning Capacities, and Costless Measures*, mimeo. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3562739

<https://mpra.ub.uni-muenchen.de/99728/>

³³ Chan, Chun Lam, Pak Hou Che, Sidhart Jaggi and Venkatesh Saligrama (2011), “Non-adaptative probabilistic group testing with noisy measurements: Near-optimal bounds with efficient algorithms”, in *49th Allerton Conference*, pp. 1832-1839.

³⁴ Baldassini, Leonardo, Oliver Johnson and Matthew Aldridge (2013), “The Capacity of Adaptive Group Testing”, *IEEE International Symposium on Information Theory*, 2676-2680.

³⁵ Aldridge, M, Johnson, O, and Scarlett, J. (2019, October 2), “Group Testing: An Information Theory Perspective,” *Foundations and Trends in Communications and Information Theory* 15(3-4): 196-392.

³⁶ Macula, Anthony J. (1997) “Error-correcting nonadaptive group testing with disjunct matrices,” *Discrete Applied Mathematics*, vol. 80, no. 2-3, pp. 217 – 222, 1997.

³⁷ Yang, Yang, Minghui Yang, Chenguang Shen, Fuxiang Wang, Jing Yuan, Jinxiu Li, Mingxia Zhang, Zhaoqin Wang, Li Xing, Jinli Wei, Ling Peng, Gary Wong, Haixia Zheng, Weibo Wu, Mingfeng Liao, Kai Feng, Jianming Li, Qianting Yang, Juanjuan Zhao, Zheng Zang Lei Liu and Yingxia Liu (2020, February). “Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections”.

³⁸ Wang, Wenling, Yanli Xu, Ruqin Gao, Roujian Lu, Kai Han, Ghuizen Wu and Wenjie Tan (2020, March). “Detection of SARS-CoV-2 in different types of clinical specimens” *JAMA*. doi:10.1001/jama.2020.3786.

³⁹ Wang, Xiong, Li Tan, Xu Wang, Weiyong Liu, Yanjun Lu, Liming Cheng, Ziyong Sun (2020), “Comparison of nasopharyngeal and oropharyngeal swabs for SARS-CoV-2 detection in 353 patients received tests with both specimens simultaneously”, *International Journal of Infectious Diseases* 94, 107–109

⁴⁰ Atia, G. and V. Saligrama, “Boolean compressed sensing and noisy group testing” *IEEE Trans. Inform. Theory*, vol. 58, no. 3, pp. 1880 –1901, March 2012.

⁴¹ Huang C, Wang Y, Li X, *et al.* (2020), “Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China”, *Lancet*.