Received: 9 August 2022 | Accepted: 24 August 2022

DOI: 10.1002/jmv.28099

SHORT COMMUNICATION

JOURNAL OF MEDICAL VIROLOGY WILEY

Estimating the undetected infections in the Monkeypox outbreak

Antonello Maruotti[1] | Dankmar Böhning[2] | Irene Rocchetti[3] | Massimo Ciccozzi[4]

[1]Department GEPLI, Libera Università Ss Maria Assunta, Rome, Italy

[2]Southampton Statistical Sciences Research Institute, University of Southampton, Southampton, UK

[3]Statistical Office—Consiglio Superiore della Magistratura, Rome, Italy

[4]Department of Medicine, Unit of Medical Statistics and Molecular Epidemiology, University Campus Bio-Medico of Rome, Rome, Italy

Correspondence
Antonello Maruotti, Department GEPLI, Libera Università Ss Maria Assunta, Via Pompeo Magno 28, Rome 00192, Italy.
Email: a.maruotti@lumsa.it

Abstract

While the number of detected Monkeypox infections are widely available, an understanding of the extent of undetected cases is urgently needed for an effective tackling of its spread. The aim of this study is to estimate the true number of Monkeypox (detected and undetected) infections in most affected countries. The question being asked is: How many cases have actually occurred? We propose a lower bound estimator for the true number of Monkeypox cases. The estimator is data-driven and can be easily computed from the cumulative distributions of weekly cases. We focused on the ratio of the total estimated cases to the observed cases on July 31, 2022: The proportion of undetected cases was relevant in all countries, with countries whose estimated *true* number of infections could be more than three times the observed one. We provided a practical contribution to the understanding of the current Monkeypox wave and reliable estimates on how many undetected cases are going around in several countries, where the epidemic spreads differently.

KEYWORDS

asymptomatic cases, capture–recapture methods, incidence indicators, Monkeypox

1 | INTRODUCTION

When an infectious disease outbreak starts, it's like a brush fire. The flames are small and, in theory, can be contained if chains of transmission are snuffed out before they spark more small fires that eventually grow into a raging epidemic. Bragazzi et al.,[1] Saxena et al.,[2] and Siddiqui et al.[3] are just a few examples of research discussing information about the recent outbreaks of human Monkeypox, epidemiology, transmission pattern, possible diagnosis techniques, therapeutics, and available preventive strategies.

The key to it all is testing,[4] a lesson that we all painfully learned early on with COVID-19 and is once again confronting as Monkeypox spreads across the world.

If that gap between test capacity and use persists, Monkeypox could spread under the radar and become endemic in several countries, despite being a much less transmissible, and thus more easily containable, virus than, for example, the SARS-COV-2.

The appearance of Monkeypox cases in many parts of the world suggests there may have been undetected transmission. But, "How many undetected cases are going around?" is still an open question we would like to give an answer. As stated by the WHO (https://www.who.int/news-room/fact-sheets/detail/monkeypox), "the extent to which asymptomatic infection may occur is unknown," and thus the number of total infections is unknown. However, the topic has been recently discussed by Ferré et al.[5] and De Baetselier et al.[6] on individual samples collected in France and Belgium. Both studies acknowledge that certain cases of Monkeypox remain undiagnosed.

TABLE 1 Estimated total cases of Monkeypox for several countries, on July 31, 2022, with confidence intervals in brackets

Country	Observed cases	Lower bound for the estimated total number of cases	Total estimate case/Observed cases
Brazil	1367	4303 (3865–4740)	3.14
Canada	676[a]	1511 (1321–1701)	2.24
France	1561[a]	4935 (4413–5456)	3.16
Germany	2591	5696 (5372–6019)	2.20
Italy	476	1024 (891–1156)	2.15
The Netherlands	876	2271 (1973–2569)	2.59
Portugal	565[a]	1029 (916–1142)	1.82
Spain	4260	13 141 (12 262–14 019)	3.08
United Kingdom	2473	7653 (6875–8433)	3.10
United States of America	5173	14 779 (14 396–15 162)	2.86

[a]Data with up to July 24, 2022.

Up to our knowledge, there are no methodological, statistical contributions to the estimation of undetected cases of the Monkeypox outbreak.

2 | METHODS

Here, we are proposing a simple and effective method to obtain reasonable point and interval estimates of the total number of Monkeypox infections in several countries. The available data just tell us a part of the story: individuals may be already infected but are not aware of it, maybe because of the absence of symptoms, or cases may be under symptomatic suspicion but the disease has not been diagnosed yet (due to the lack of testing capacity). The total number of cases is thus unknown, and general comments on the spread of the epidemic are thus partial as based on a fraction of the total cases. It is important to mention that in the first months of the Monkeypox outbreak, no mass or systematic testing of the population had taken place; it is very likely that the *true* number of infections is much higher than the observed one.

In detail, we introduce an estimator based on a capture–recapture (CR) approach. The CR method should be considered as the gold standard for counting when it is impossible to identify each case and large undercounts will occur.[7] CR methods were originally developed in the ecological setting with the aim of estimating the unknown size of a (possibly elusive) population and then they started to be applied also to epidemiological and health sectors (see Böhning et al.[8]).

We will denote with $N(t)$ the cumulative count of infections at week t where $t = t_0, \ldots, t_m$. Hence, $\Delta N(t) = N(t) - N(t-1)$ are the number of new infections at week t where $t = t_0 + 1, \ldots, t_m$. Our primary outcome is the number of weekly new cases of Monkeypox as reported by official sources.[9] We look at the most 10 affected countries worldwide, collected from May 23, 2022 until July 31, 2022.

We apply a lower-bound Chao estimator[10] weekly wise: The estimated number of undetected cases is given by the ratio between the square of the frequency of those identified exactly once and of those identified exactly twice. Here, $\Delta N(t)$ corresponds to the infected people identified just once, and $\Delta N(t-1)$ is the number of those identified twice. To avoid bias estimates, we modify the lower bound Chao estimator as in Böhning et al.,[11] such that we are able to give the estimate for the number of hidden infections as

$$H_{t_0} = \sum_{t=t_0+1}^{t_m} \frac{\Delta N(t)[\Delta N(t) - 1]}{1 + \Delta N(t-1)}. \quad (1)$$

The final estimate of the total size of infection is then given as what has been observed at the end of the observational window t_m and the estimate of the hidden numbers:

$$\text{total size of infections} = N(t_m) + H_{t_0}. \quad (2)$$

The 95% confidence intervals are given by (see, e.g., Niwitpong et al.[12])

$$H_{t_0} \pm 1.96\sqrt{\sum_{t=t_0+1}^{t_m} \hat{\text{Var}}[H(t)]} \quad (3)$$

with

$$\hat{\text{Var}}[H(t)] = \frac{[\Delta N(t)]^4}{[1 + \Delta N(t-1)]^3} + \frac{4[\Delta N(t)]^3}{[1 + \Delta N(t-1)]^2} + \frac{[\Delta N(t)]^2}{[1 + \Delta N(t-1)]}. \quad (4)$$

3 | RESULTS

The sum of the *dark number*, that is, the estimated number of undetected cases, and the observed cases is displayed in Table 1 for each country, along with the 95% confidence interval (providing a measure of uncertainty of the estimate) and the ratio between the

estimated total cases and the observed ones, as a measure of how severe is the impact on undetected cases on the spread of the Monkeypox.

The obtained estimates are clear-cut: The proportion of undetected cases is relevant in all countries. There is, however, some heterogeneity across countries on the impact of missed cases on the total number of cases. In Brazil, France, Spain, and United Kingdom, the number of infections could reasonably be more than three times the observed number, and this is also reflected by recent sharp increases in the number of observed cases. In Canada, Germany, and Italy, the estimated number of cases is a bit above two times the observed ones, with the Netherlands and the United States of America lying in between, even though the situation in the United States of America is likely to be more at risk. At last, Portugal shows an estimated ratio below 2, that is, the missing infection phenomenon is still rather limited.

4 | DISCUSSION

The proposed method answers to a fundamental open question: "How many undetected cases are going around?" We remark that lower bound of the number of total infections is provided, but this information may be treated as a starting point whenever interventions and tools to dampen the spread of the epidemic are rolled out. This is a relevant result as it provides reasonable information to the policymakers about the undetected cases and the magnitude of this phenomenon may have at least, so that national health systems may be aware of the minimum number of cases that may demand health care services.

The sudden appearance of Monkeypox in multiple countries across the world indicates the virus has been spreading undetected for some time outside the West and Central African nations where it is usually found. Having an estimate of this phenomenon is fundamental to apply nonpharmaceutical interventions to contain the spread of the virus. And when a virus spreads cryptically like this, it can be really hard to stop, and there's a chance it could become a long-term problem.

Further attention should be in place as cases may be undetected because the disease looks different than what's described in medical textbooks. The concern on how the virus might change, especially in terms of how it spreads, is real: there's a possibility it has become more contagious.

We believe it is not too late to contain Monkeypox. However, to prevent onward spread, as a general guide, high-quality data are required. Indeed, when the poor quality of the data, as often happens at the beginning of an outbreak, does not allow to correctly apply sophisticated models, a robust data-driven approach, like the CR approach here proposed, could be used as a starting point for any analyses.

Contact tracing and isolating patients who have Monkeypox are crucial to stopping the spread and are still the only tools we have at the moment to manage the epidemic, along with strengthening national surveillance.[13] Vaccination should be recommended for all contacts of positives and for at risk subjects.[14]

The proposed approach is simple and can be straightforwardly applied by nonexperts. Being a nonparametric approach, we avoid the specification of a distribution for the observed counts, which may be restrictive at the beginning of the outbreak where the number of observed weeks is very limited. Indeed, the proposed estimator can be computed directly from the cumulative distribution of the infected counts. Moreover, the code to obtain the estimates is available in the *asymptor* package[15] of the R software.

As it stands, the proposed estimator accounts neither for the number of tested people nor for the number of weekly recovered people. Testing capacity, as well as other external variables, could be straightforwardly included as discussed in Böhning et al.[16] or Dotto and Farcomeni,[17] with minor efforts. Similarly, to account for the number of recovered, let us denote with $D(t)$ the cumulative count of recovered at week t where $t = t_0, \ldots, t_m$. Hence, $\Delta D(t) = D(t) - D(t-1)$ is the number of new recovered at week t where $t = t_0 + 1, \ldots, t_m$. The proposed estimator is the given by

$$H_{t_0} = \sum_{t=t_0+1}^{t_m} \frac{\Delta N(t)[\Delta N(t) - 1]}{1 + \Delta N(t-1) - \Delta D(t)}, \quad (5)$$

where $\Delta N(t-1) - \Delta D(t)$ is set to 0 if it becomes negative. The variance should be then modified accordingly. Unfortunately, none of these two information, the number of tested and recovered people, is currently available.

ACKNOWLEDGMENTS

Open Access Funding provided by Libera Universita Maria Santissima Assunta within the CRUI-CARE Agreement.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data are freely available at https://ourworldindata.org/monkeypox

ORCID

Antonello Maruotti http://orcid.org/0000-0001-8377-9950
Massimo Ciccozzi http://orcid.org/0000-0003-3866-9239

REFERENCES

1. Bragazzi NL, Kong JD, Mahroum N, et al. Epidemiological trends and clinical features of the ongoing monkeypox epidemic: a preliminary pooled data analysis and literature review. *J Med Virol*. 2022. doi:10.1002/jmv.27931
2. Saxena SK, Ansari S, Maurya VK, et al. Re-emerging human monkeypox: a major public-health debacle. *J Med Virol*. doi:10.1002/jmv.27902
3. Siddiqui MO, Syed MA, Tariq R, Mansoor S. Multicounty outbreak of monkeypox virus—challenges and recommendations. *J Med Virol*. 2022. doi:10.1002/jmv.27966
4. Suran M. Expanding Monkeypox testing. *JAMA*. 2022;328:321.
5. Ferré VM, Bachelard A, Zaidi M, et al. Detection of Monkeypox virus in anorectal swabs from asymptomatic men who have sex with men

in a sexually transmitted infection screening program in Paris, France. *Ann Intern Med*. 2022. doi:10.7326/M22-2183
6. De Baetselier I, Van Dijck C, Kenyon C, et al. Retrospective detection of asymptomatic monkeypox virus infections among male sexual health clinic attendees in Belgium. *Nat Med*. 2022. doi:10.1038/s41591-022-02004-w
7. Lange JH, LaPorte RE. Capture–recapture method should be used to count how many cases of SARS really exist. *BMJ*. 2003;326:1396.
8. Böhning D, van der Heijden PGM, Bunge J. *Capture–Recapture Methods for the Social and Medical Science*. CRC Press; 2019.
9. Mathieu E, Spooner F, Dattani S, Ritchie H, Roser M. "Monkeypox". Accessed August 8, 2022. https://ourworldindata.org/monkeypox
10. Chao A, Colwell RK. Thirty years of progeny from Chao's inequality: estimating and comparing richness with incidence data and incomplete sampling. *SORT Stat Oper Res Trans*. 2017;41:3-54.
11. Böhning D, Rocchetti I, Maruotti A, Holling H. Estimating the undetected infections in the Covid-19 outbreak by harnessing capture–recapture methods. *Int J Infect Dis*. 2020;97:197-201.
12. Niwitpong SA, Boehning D, van der Heijden PG, Holling H. Capture–recapture estimation based upon the geometric distribution allowing for heterogeneity. *Metrika*. 2013;76:495-519.
13. Nuzzo JB, Borio LL, Gostin LO. The WHO declaration of monkeypox as a global public health emergency. *JAMA*. 2022;328:615-617. doi:10.1001/jama.2022.12513
14. Guarner J, Del Rio C, Malani PN. Monkeypox in 2022—what clinicians need to know. *JAMA*. 2022;328:139-140.
15. Gruson H. Asymptor: estimate the lower and upper bound of asymptomatic cases in an epidemic using the capture/recapture Methods. Accessed August 8, 2022. https://CRAN.R-project.org/package=asymptor
16. Böhning D, Vidal-Diez A, Lerdsuwansri R, Viwatwongkasem C, Arnold M. A generalization of Chao's estimator for covariate information. *Biometrics*. 2013;69(4):1033-1042.
17. Dotto F, Farcomeni A. A generalized Chao estimator with measurement error and external information. *Environ Ecol Stat*. 2018;25(1):53-69.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Maruotti A, Böhning D, Rocchetti I, Ciccozzi M. Estimating the undetected infections in the Monkeypox outbreak. *J Med Virol*. 2022;1-4. doi:10.1002/jmv.28099