

Sussex Research

The prevalence of onchocerciasis in Africa and Yemen, 2000–2018: a geospatial analysis

Chris A Schmidt, Elizabeth A Cromwell, Elex Hill, Katie M Donkers, Megan F Schipp, Kimberly B Johnson, David M Pigott, Jaffar Abbas, Victor Adekanmbi, Olatunji O Adetokunboh, Kebede Kassaye, others

Publication date

07-06-2023

Licence

This work is made available under the [CC](https://creativecommons.org/licenses/by/4.0/) BY 4.0 licence and should only be used in accordance with that licence. For more information on the specific terms, consult the repository record for this item.

Document Version

Published version

Citation for this work (American Psychological Association 7th edition)

Schmidt, C. A., Cromwell, E. A., Hill, E., Donkers, K. M., Schipp, M. F., Johnson, K. B., Pigott, D. M., Abbas, J., Adekanmbi, V., Adetokunboh, O. O., Kassaye, K., & others, . (2022). The prevalence of onchocerciasis in Africa and Yemen, 2000–2018: a geospatial analysis (Version 1). University of Sussex. https://hdl.handle.net/10779/uos.23308874.v1

Published in

BMC Medicine

Link to external publisher version

<https://doi.org/10.1186/s12916-022-02486-y>

Copyright and reuse:

This work was downloaded from Sussex Research Open (SRO). This document is made available in line with publisher policy and may differ from the published version. Please cite the published version where possible. Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners unless otherwise stated. For more information on this work, SRO or to report an issue, you can contact the repository administrators at [sro@sussex.ac.uk.](mailto:sro@sussex.ac.uk) Discover more of the University's research at <https://sussex.figshare.com/>

List of Supplementary Figures

65 **1.0 GATHER compliance**

66 **Supplementary Table 1. Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER)**

67 **checklist.** $\begin{array}{c} 66 \\ 67 \\ 68 \end{array}$

70 **2.0 Supplementary discussion**

71 This document outlines the major data processing, modelling, and validation steps for the onchocerciasis prevalence
72 analysis described in the main text (Supplementary Figure 1). We present a detailed description of m

-
- 22 analysis described in the main text (Supplementary Figure 1). We present a detailed description of model inputs,

23 including data coverage, covariate sources, and geo-referencing. The geospatial model is described alo including data coverage, covariate sources, and geo-referencing. The geospatial model is described along with
- model validation metrics.
- 75

76 77 **Supplementary Figure 1. Flowchart of major steps in data processing and modelling of onchocerciasis** 78 **prevalence.**

79 **3.0 Supplementary data**

80 In the following section, we present a detailed summary of the data inputs used to estimate the prevalence of onchocerciasis in Africa and Yemen. Broadly, we aimed to include all published sources of onchocerciasis in

81 onchocerciasis in Africa and Yemen. Broadly, we aimed to include all published sources of onchocerciasis infection
82 percylence, as well as routine programme monitoring data collected to monitor progress toward onchoce

- prevalence, as well as routine programme monitoring data collected to monitor progress toward onchocerciasis
- elimination. Data inputs were retained for analysis if they could be accurately geo-referenced. $\frac{83}{83}$
-

85 **3.1 Geographical restrictions**

- 86 Supplementary Table 2 lists all countries included in our MBG (model-based geostatistical) modelling region.
87 Inclusion was partially based on ESPEN (Expanded Special Project for Elimination of Neglected Tropical Dis
- 87 Inclusion was partially based on ESPEN (Expanded Special Project for Elimination of Neglected Tropical Diseases)
- 88 onchocerciasis endemicity classifications,(1) with extensions to some neighbouring countries outside the WHO
89 (World Health Organization) AFRO (Regional Office for Africa) region which have evidence of onchocerciasis
- 89 (World Health Organization) AFRO (Regional Office for Africa) region which have evidence of onchocerciasis
90 endemicity or uncertain endemicity status (Sudan, Somalia, and Yemen), Portions of Mauritius, Namibia, and
- 90 endemicity or uncertain endemicity status (Sudan, Somalia, and Yemen). Portions of Mauritius, Namibia, and
91 Zambia are considered by ESPEN as candidates for future elimination mapping but were not included in our
- 91 Zambia are considered by ESPEN as candidates for future elimination mapping but were not included in our
92 geospatial model. Converselv, The Gambia is considered non-endemic but we included it in our model for
- 92 geospatial model. Conversely, The Gambia is considered non-endemic but we included it in our model for
- 93 geographical continuity. Countries in the Americas with historical or residual onchocerciasis burdens were not
94 modelled due to their highly localised endemicity.
- modelled due to their highly localised endemicity.

95 **Supplementary Table 2. Geographical restrictions.**

96 The geographical definition of the modelling region is indicated, with country-level onchocerciasis endemicity
97 status per ESPEN, and which countries were included in our geospatial analysis from 2000 to 2018. Endemic

- 97 status per ESPEN, and which countries were included in our geospatial analysis from 2000 to 2018. Endemicity
98 classifications were derived from ESPEN data at the level of intervention units (IU) (retrieved 14 February
- 98 classifications were derived from ESPEN data at the level of intervention units (IU) (retrieved 14 February 2020).
99 Countries outside the WHO AFRO region are not covered by the scope of ESPEN; countries in the WHO EMR
- 99 Countries outside the WHO AFRO region are not covered by the scope of ESPEN; countries in the WHO EMRO
100 (Eastern Mediterranean Regional Office) region are listed here if they are in Africa or have evidence of
- 100 (Eastern Mediterranean Regional Office) region are listed here if they are in Africa or have evidence of onchocerciasis endemicity or uncertain status. We considered countries endemic if they had at least one
- 101 onchocerciasis endemicity or uncertain status. We considered countries endemic if they had at least one intervention
102 unit that was flagged as endemic by ESPEN. Number of observations: The number of data rows in the
- unit that was flagged as endemic by ESPEN. Number of observations: The number of data rows in the final dataset 103 for a given country.
- 104
	- **Location ESPEN endemicity Included in 2000–2018 model? Number of Observations** *AFRO*(1) Algeria and the Non-endemic contract the no no contract of the Non-endemic contract to the no contract of the Non-endemic contract in the no contract of the Non-endemic contract in the normal contract of the Non-endemic co Angola Endemic yes 876 Benin and Endemic Contract the Endemic Solution of the Solutio Botswana and the Southern Non-endemic the state of the no -Burkina Faso Endemic yes 72 Burundi Endemic yes 263 Cabo Verde **Non-endemic** Non-endemic no no -Cameroon **Endemic** Endemic yes 1216 Central African Republic Endemic yes 1079 Chad Endemic yes 753 Comoros Non-endemic no no -Republic of the Congo **Endemic** Endemic yes 490 Côte d'Ivoire Endemic yes 95 Democratic Republic of the Congo Endemic and the set of t Equatorial Guinea Endemic yes 341 Eritrea and the South Mon-endemic contract the contract of the no -Eswatini Non-endemic no no -

106

107 **3.2 Systematic review**

108 Articles related to onchocerciasis were found by searching PubMed, Scopus, and Web of Science using the 109 following keywords: "Oncho", "river blindness", "O. Volvulus", "robles disease", "blinding filariasis", "coa

109 following keywords: "Oncho", "river blindness", "O. Volvulus", "robles disease", "blinding filariasis", "coast erysipelas", and "sowda". A systematic review of these reports, all published before July 7, 2017, was then

110 erysipelas", and "sowda". A systematic review of these reports, all published before July 7, 2017, was then
111 conducted. A second round of formal reviews was undertaken on June 6, 2019, to cover new articles publis

111 conducted. A second round of formal reviews was undertaken on June 6, 2019, to cover new articles published after 112 July 7, 2017, and followed the same process as the first round.

July 7, 2017, and followed the same process as the first round. 113

114 **3.2.1 Systematic review data processing**
115 The systematic review process is illustrated

115 The systematic review process is illustrated in Supplementary Figure 2. Throughout the systematic review we
116 excluded publications that met the following criteria: no measurement of onchocerciasis prevalence, data c 116 excluded publications that met the following criteria: no measurement of onchocerciasis prevalence, data collected 117 before 1985, case-control studies, qualitative research publications, duplicative data from cohort 117 before 1985, case-control studies, qualitative research publications, duplicative data from cohort studies, hospital-
118 based studies, and publications that did not report the location of data collection. The search 118 based studies, and publications that did not report the location of data collection. The search identified 4048
119 bublications, which were reduced to 579 after screening titles and abstracts. A full-text review was c 119 publications, which were reduced to 579 after screening titles and abstracts. A full-text review was completed for the 120 remaining publications. The full-text review yielded 259 publications which met the inclusion c 120 remaining publications. The full-text review yielded 259 publications which met the inclusion criteria and were
121 extracted. The literature review was updated on June 6, 2019, by searching PubMed with the same search 121 extracted. The literature review was updated on June 6, 2019, by searching PubMed with the same search string used
122 in 2017 for articles published after July 7, 2017. The search returned 304 results, which were narr 122 in 2017 for articles published after July 7, 2017. The search returned 304 results, which were narrowed to 115
123 articles after screening titles and abstracts. A full-text review resulted in 22 articles that were eli 123 articles after screening titles and abstracts. A full-text review resulted in 22 articles that were eligible for extraction.
124 Additional publications were identified outside of the literature review and screened for

124 Additional publications were identified outside of the literature review and screened for inclusion by the same
125 eriteria, up through April 14, 2020. The final dataset drew from 290 articles. Overall, 16 096 datapoi

- 125 criteria, up through April 14, 2020. The final dataset drew from 290 articles. Overall, 16 096 datapoints were
126 extracted and geo-located from 290 publications. Data extracted from each source included survey year,
- 126 extracted and geo-located from 290 publications. Data extracted from each source included survey year, age-range of individuals tested, diagnostic information, sample size, number of individuals tested positive, and sa 127 of individuals tested, diagnostic information, sample size, number of individuals tested positive, and sampling strategy details. strategy details.

129

$\frac{130}{131}$

131 **Supplementary Figure 2. Onchocerciasis article review and data extraction flowchart.**

132 Each step of the extraction process is outlined, from article identification and screening to extraction, including the number of articles or records that were processed or removed in each step before reaching the fina

133 number of articles or records that were processed or removed in each step before reaching the final dataset.
134 dditional articles outside of the literature review were identified and screened for inclusion ("ongoing

134 Additional articles outside of the literature review were identified and screened for inclusion ("ongoing process") on an ongoing basis up until April 14, 2020.

- 135 an ongoing basis up until April 14, 2020.
- 136

137 **3.3 Geo-positioning**

138 Geographical information associated with onchocerciasis prevalence data were verified and geo-located to ensure 139 accuracy. Data associated with locations smaller than 5×5 km were treated as points and geo-locat

accuracy. Data associated with locations smaller than 5×5 km were treated as points and geo-located as

140 latitude/longitude coordinates. Coordinates provided by a data source were mapped to ensure that the coordinates were located in the correct administrative units. If coordinates were not reported, points were geolocate

- 141 were located in the correct administrative units. If coordinates were not reported, points were geolocated and vetted 142 by comparing results from Google Maps, Fuzzy Gazetteer, Geonames, and Open Street Map. Locations
- 142 by comparing results from Google Maps, Fuzzy Gazetteer, Geonames, and Open Street Map. Locations larger than 5
143 \times 5 km were treated as polygon data and geo-located to the most granular administrative boundary po

¹⁴³ \times 5 km were treated as polygon data and geo-located to the most granular administrative boundary possible (most commonly district level). Custom shapefiles were created to geolocate areas that did not align with a

- 144 commonly district level). Custom shapefiles were created to geolocate areas that did not align with administrative boundaries. In the event that a literature source only included a map of locations sampled without any
- 145 boundaries. In the event that a literature source only included a map of locations sampled without any other
146 information, ArcGIS software was utilised to overlay the map onto existing administrative boundaries, and 146 information, ArcGIS software was utilised to overlay the map onto existing administrative boundaries, and location coordinates or custom polygons were manually created and recorded. Prevalence data that were reported f
- 147 coordinates or custom polygons were manually created and recorded. Prevalence data that were reported for
148 dministrative areas were matched to their appropriate polygon by searching our administrative shapefile dat
- 148 administrative areas were matched to their appropriate polygon by searching our administrative shapefile database.
149 If place names were unidentifiable across multiple shapefile libraries or geo-referencing sources,
- 149 If place names were unidentifiable across multiple shapefile libraries or geo-referencing sources, they were excluded from the analysis.
- from the analysis.
- 151

152 **3.4 Data processing**

- 153 Data were excluded from the analysis for the following reasons: did not report survey year (N = 8236); did not report post-1988 data (N = 2648); did not report skin snip microscopy or nodule prevalence (N = 12 121); d
- 154 report post-1988 data (N = 2648); did not report skin snip microscopy or nodule prevalence (N = 12 121); did not report sample size or prevalence (N = 30); or did not report data for countries in the modelling region 155 report sample size or prevalence ($N = 30$); or did not report data for countries in the modelling region or could not be accurately georeferenced ($N = 580$). Duplicate records were identified and excluded ($N = 1592$);
- 156 be accurately georeferenced (N = 580). Duplicate records were identified and excluded (N = 1592); these generally reflected different data sources (eg, literature extraction and ESPEN) which contained the same survey,
- 157 reflected different data sources (eg, literature extraction and ESPEN) which contained the same survey, as judged by
- 158 identical year, location, sample size, cases, and reference when available, although there was some deduplication within individual data sources. Prior to polygon resampling (see section 5.2), the dataset consisted of
- 159 within individual data sources. Prior to polygon resampling (see section 5.2), the dataset consisted of 18 116
160 georeferenced prevalence observations, consisting of 17 896 point-referenced observations and 220 areal
- 160 georeferenced prevalence observations, consisting of 17 896 point-referenced observations and 220 areal
- 161 observations. A total of 14 314 observations represented prevalence based on nodule examinations, and 3802
- 162 observations represented microfiladermia prevalence as measured by skin snip assays. After resampling, the full
- 163 modelling dataset consisted of 20 124 georeferenced datapoints.
- 164
165
- 165 Supplementary Table 3 provides citations for data sources used in our onchocerciasis MBG model. The
- 166 geographical coverage of the final dataset is summarised in Supplementary Figure 3.

167 **Supplementary Table 3. Citations for data inputs.**

168 The NID is a unique identifier cataloguing all data inputs in the Global Health Data Exchange 168
169
170
171

169 (http://ghdx.healthdata.org). Note: Records are listed here in alphabetical order by geography, but some sources

provided data for multiple countries; such sources are listed here only once.

NID	Geographies	Citation
332798	Angola	Carme B, Ntsoumou-Madzou V, Samba Y, Yebakima A. Prevalence of depigmentation of the shins: a simple and cheap way to screen for severe endemic onchocerciasis in Africa. Bull World Health Organ. 1993; 71(6): 755-8.
334477	Benin	Gallin M, Adams A, Kruppa TF, Gbaguidi EA, Massougbodji A, Sadeler BC, Brattig N, Erttmann KD. Epidemiological studies of onchocerciasis in southern Benin. Trop Med Parasitol. 1993; 44(2): 69-74.
125405	Burkina Faso	Kabore JK, Cabore JW, Melaku Z, Druet-Cabanac M, Preux PM. Epilepsy in a focus of onchocerciasis in Burkina Faso. Lancet. 1996; 347(9004): 836.
334475	Burkina Faso	De Sole G, Remme J. Importance of migrants infected with Onchocerca volvulus in west African river valleys protected by 14 to 15 years of Simulium control. Trop Med Parasitol. 1991; 42(2): 75-8.
332903	Burkina Faso, Côte d'Ivoire	Toè L, Adjami AG, Boatin BA, Back C, Alley ES, Dembélé N, Brika PG, Pearlman E, Unnasch TR. Topical application of diethylcarbamazine to detect onchocerciasis recrudescence in west Africa. Trans R Soc Trop Med Hyg. 2000; 94(5): 519-25.
334481	Burundi	Newell E d. Comparison of the use of skin scarification and skin biopsies to determine the prevalence and intensity of Onchocerca volvulus infection. Ann Trop Med Parasitol. 1997; 91(6): 633.
338571	Burundi	Newell ED, Hicuburundi B, Ndimuruvugo N. [Endemicity and clinical manifestations of onchocerciasis in the province of Bururi, Burundi]. Trop Med Int Health. 1997; 2(3): 218–26.
338573	Burundi	Newell ED, Ndimuruvugo N, Nimpa D. [Endemicity and clinical manifestations of onchocerciasis in the provinces of Cibitoke and Bubanza (Burundi)]. Bull Soc Pathol Exot. 1997; 90(5): 353-7.
136492	Cameroon	Cho-Ngwa F, Amambua AN, Ambele MA, Titanji VPK. Evidence for the exacerbation of lymphedema of geochemical origin, podoconiosis, by onchocerciasis. J Infect Public Health. 2009; 2(4): 198-203.
159316	Cameroon	Matthews GA, Dobson HM, Nkot PB, Wiles TL, Birchmore M. Preliminary examination of integrated vector management in a tropical rainforest area of Cameroon. Trans R Soc Trop Med Hyg. 2009; 103(11): 1098-104.
324729	Cameroon	Kamga G-R, Dissak-Delon FN, Nana-Djeunga HC, Biholong BD, Mbigha- Ghogomu S, Souopgui J, Zoure HGM, Boussinesq M, Kamgno J, Robert A. Still mesoendemic onchocerciasis in two Cameroonian community-directed treatment with ivermectin projects despite more than 15 years of mass treatment. Parasites Vectors. 2016; $9(1)$: 581.
327960	Cameroon	Wanji S, Kengne-Ouafo JA, Esum ME, Chounna PWN, Tendongfor N, Adzemye BF, Eyong JEE, Jato I, Datchoua-Poutcheu FR, Kah E, Enyong P, Taylor DW. Situation analysis of parasitological and entomological indices of onchocerciasis transmission in three drainage basins of the rain forest of South West Cameroon after a decade of ivermectin treatment. Parasites Vectors. 2015; 8: 202.
327968	Cameroon	Katabarwa MN, Eyamba A, Chouaibou M, Enyong P, Kuété T, Yaya S, Yougouda A, Baldiagaï J, Madi K, Andze GO, Richards F. Does onchocerciasis transmission take place in hypoendemic areas? a study from the North Region of Cameroon. Trop Med Int Health. 2010; 15(5): 645-52.

172

$\frac{175}{176}$

Supplementary Figure 3. Africa and Yemen data coverage maps.

177 Locations of the unique point and polygon data used in modelling, grouped by years: 1988–1999, 2000–2004, 2005– 177 Locations of the unique point and polygon data used in modelling, grouped by years: 1988–1999, 2000–2004, 200
178 2009, and 2010–2018. Hatched countries were not included in the MBG modelling region. Data that were abl

178 2009, and 2010–2018. Hatched countries were not included in the MBG modelling region. Data that were able to be

blue polygons.

181 **3.5 APOC survey year matching** 182 The African Programme for Onchocerciasis Control (APOC), (3) which operated from 1995 to 2015, relied on Rapid
183 Epidemiological Mapping of Onchocerciasis (REMO) to map baseline onchocerciasis prevalence and endemici 183 Epidemiological Mapping of Onchocerciasis (REMO) to map baseline onchocerciasis prevalence and endemicity in 184 19 onchocerciasis-endemic African countries outside the Onchocerciasis Control Programme (OCP) focal region 185 (APOC was later expanded to additional countries). REMO(4.5) mapping, involving village-level surveys of n 185 (APOC was later expanded to additional countries). REMO(4,5) mapping, involving village-level surveys of nodule
186 (onchocercoma) prevalence among adult males, was completed for 12 APOC countries by 2001.(6) We obtain 186 (onchocercoma) prevalence among adult males, was completed for 12 APOC countries by 2001.(6) We obtained
187 these pre-2001 nodule mapping data, when available, from the ESPEN data portal, although detailed information 187 these pre-2001 nodule mapping data, when available, from the ESPEN data portal, although detailed information on survey years were lacking for many $(N = 4409)$ of these datapoints. So that these essential baseline mapp 188 survey years were lacking for many $(N = 4409)$ of these datapoints. So that these essential baseline mapping surveys could inform our spatiotemporal model of onchocerciasis prevalence, we investigated the pre-2001 REMO 189 could inform our spatiotemporal model of onchocerciasis prevalence, we investigated the pre-2001 REMO mapping
190 surveys in APOC countries and attempted to narrow the possible range of survey vears for these data. As 190 surveys in APOC countries and attempted to narrow the possible range of survey years for these data. As widespread 191 use of REMO by APOC began in 1996.(6) and these data were indicated by the ESPEN database as repres 191 use of REMO by APOC began in 1996,(6) and these data were indicated by the ESPEN database as representing
192 pre-2001 surveys, we considered 1996 and 2000 as the initial lower and upper bounds for survey years. We nex 192 pre-2001 surveys, we considered 1996 and 2000 as the initial lower and upper bounds for survey years. We next
193 identified survey locations for which the first year of onchocerciasis MDA (mass drug administration) oc 193 identified survey locations for which the first year of onchocerciasis MDA (mass drug administration) occurred prior
194 to 2001 (see section 4.2 below) and assumed that baseline mapping occurred at that location no la 194 to 2001 (see section 4.2 below) and assumed that baseline mapping occurred at that location no later than the
195 ereceding year, treating that year as the new upper bound. Finally, we examined the available literature 195 preceding year, treating that year as the new upper bound. Finally, we examined the available literature on pre-2001 196 APOC REMO surveys, by country, to further narrow the possible range of survey years. Our decision 196 APOC REMO surveys, by country, to further narrow the possible range of survey years. Our decisions around
197 urvey year estimates are summarised in Supplementary Table 4. For each datapoint, we specified the final sur 197 survey year estimates are summarised in Supplementary Table 4. For each datapoint, we specified the final survey
198 vear as the median of the lower and upper year bounds (rounding down), considering this approach to b 198 year as the median of the lower and upper year bounds (rounding down), considering this approach to balance the potential conflicting biases from incorrect survey years on the temporal trends in the MBG model and the 199 potential conflicting biases from incorrect survey years on the temporal trends in the MBG model and the association with environmental and sociodemographic covariates in child models. association with environmental and sociodemographic covariates in child models.

 $\frac{201}{202}$ 202 Survey year data were also missing for some Phase 1a and 1b monitoring and evaluation surveys in the APOC data
203 etrieved from ESPEN. We identified survey years for most of these data by matching them to the location 203 retrieved from ESPEN. We identified survey years for most of these data by matching them to the locations, sample
204 sizes, cases, and number of villages reported by Tekle et al.,(7) supplemented with additional sourc 204 sizes, cases, and number of villages reported by Tekle et al.,(7) supplemented with additional sources(8,9) for some
205 countries. We identified survey years for a total of 1001 observations. Survey years could not be 205 countries. We identified survey years for a total of 1001 observations. Survey years could not be reliably identified
206 for some Phase 1a and 1b data, specifically Phase 1a data for Kasai Occidental (Democratic Repub 206 for some Phase 1a and 1b data, specifically Phase 1a data for Kasai Occidental (Democratic Republic of the Congo, 207 DRC), Phase 1a and 1b data for Bioko Island (Equatorial Guinea), Phase 1a data for Kogi (Nigeria), a 207 DRC), Phase 1a and 1b data for Bioko Island (Equatorial Guinea), Phase 1a data for Kogi (Nigeria), and some Phase 1b data for Uganda; these data were excluded from our analysis. We assumed sample age ranges of 5–99 yea 208 1b data for Uganda; these data were excluded from our analysis. We assumed sample age ranges of 5–99 years of 209 age for Phase 1a and 1b data(7) unless the ESPEN dataset indicated otherwise. age for Phase 1a and 1b data(7) unless the ESPEN dataset indicated otherwise.

210 **Supplementary Table 4. APOC REMO survey year investigations.** CDTI: Community-directed treatment with ivermectin.

213 **4.0 Supplementary covariates**

214 **4.1 Pre-existing covariates considered for analysis**

A variety of environmental and sociodemographic variables were used to predict all-age onchocerciasis prevalence.

216 Where available, the finest spatiotemporal resolution of gridded datasets was used. Geospatial covariate rasters were

217 resampled to \sim 5 km in GeoTiff format, for consistent modelling. Where data coverage was inconsistent between our

218 standard mask and the input data, either a local average or nearest neighbour method (depending on data type) was
219 used to fill spatial data gaps. Data from the nearest vear available were used if covariate coverage

- used to fill spatial data gaps. Data from the nearest year available were used if covariate coverage did not include all
- 220 model years. Supplementary Table 5 contains a full list of covariates considered in our analysis.

221 **4.2 Creation of MDA covariate**

222 Data used in the creation of the onchocerciasis MDA covariate were downloaded from the ESPEN Portal in April
223 2019.(19) The ESPEN data included the MDA start year and cumulative number of MDA rounds for each IU. 223 2019.(19) The ESPEN data included the MDA start year and cumulative number of MDA rounds for each IU.
224 ArcGIS software was then used to join the MDA data to the NTD (Neglected Tropical Disease) Implementation 224 ArcGIS software was then used to join the MDA data to the NTD (Neglected Tropical Disease) Implementation
225 Unit Shapefile maintained by the Task Force for Global Health. A custom polygon was created for Sudan by 225 Unit Shapefile maintained by the Task Force for Global Health. A custom polygon was created for Sudan by
226 referencing annual reports from The Carter Center. The MDA covariate value was set calculated as the cumul 226 referencing annual reports from The Carter Center. The MDA covariate value was set calculated as the cumulative number of rounds (defined as total number of rounds implemented) for the IU, and then converted to a raste 227 number of rounds (defined as total number of rounds implemented) for the IU, and then converted to a raster for use in geospatial analysis. Data for lymphatic filariasis MDA were obtained from WHO(20) for the years du 228 in geospatial analysis. Data for lymphatic filariasis MDA were obtained from WHO(20) for the years during which
229 EF (lymphatic filariasis) MDA was conducted, by IU, and joined onto the IU shapefile maintained by ESP 229 LF (lymphatic filariasis) MDA was conducted, by IU, and joined onto the IU shapefile maintained by ESPEN. So that the effects of MDA for onchocerciasis and lymphatic filariasis were not double-counted in locations wher 230 that the effects of MDA for onchocerciasis and lymphatic filariasis were not double-counted in locations where both were indicated, a composite MDA covariate was produced which indicated the cumulative number of years were indicated, a composite MDA covariate was produced which indicated the cumulative number of years for 232 which either onchocerciasis or lymphatic filariasis MDA was indicated. This composite covariate was used in modelling. modelling.

234

235 **4.3 Covariate reduction**

 236 High collinearity among covariates may lead to unstable model coefficients and unreliable predictions.³⁷ To reduce

- 237 this problem, we derived a reduced covariate set using analysis of variance inflation factors (VIF).(22) Starting with
238 the full list of covariates, we iteratively removed covariates with the highest VIF values unti
- 238 the full list of covariates, we iteratively removed covariates with the highest VIF values until all remaining
239 covariates had a VIF below 3.0. The reduced covariate set was used in fitting the MBG model and for
- 239 covariates had a VIF below 3.0. The reduced covariate set was used in fitting the MBG model and for spatiotemporal predictions. Supplementary Table 5 indicates the covariates that were retained in the fil
- spatiotemporal predictions. Supplementary Table 5 indicates the covariates that were retained in the final model,
- 241 with representative plots provided for each covariate in Supplementary Figure 4. All variables were matched to the
- 242 year of reported prevalence data, without any temporal lags (eg, temperature values for the year 2000 were joined to
- prevalence data for the year 2000).

244 **Supplementary Table 5. Covariates considered or retained for modelling, 1988–2018.**

245 All covariates considered for inclusion as predictors of onchocerciasis prevalence. Short names are provided for reference in later figures and tables. Covariates retained after VIF analysis are indicated with a (+).

retained after VIF analysis are indicated with a $(+)$.

247

251 **Supplementary Figure 4. Africa and Yemen covariate values.**

252 15 environmental or sociodemographic variables were used as inputs for the modelling process for Africa and
253 Yemen (river size is not shown in this figure). Time-varying covariates are presented here for the year 20 253 Yemen (river size is not shown in this figure). Time-varying covariates are presented here for the year 2015. Please refer to Supplementary Table 5 for the corresponding citations for each covariate.

refer to Supplementary Table 5 for the corresponding citations for each covariate.

255 **5.0 Supplementary methods**

- 256 **5.1 Age and diagnostic crosswalks** 257 Surveys of onchocerciasis prevalence have varied in both their sampled age ranges and in the diagnostic tests they
258 utilised. For example, epidemiological surveys during the OCP program relied on skin snip (microfil
- 258 utilised. For example, epidemiological surveys during the OCP program relied on skin snip (microfiladermia)
259 examinations of individuals aged 5 years or older. (23.24) while APOC relied on the REMO (Rapid Epidemiol
- 259 examinations of individuals aged 5 years or older, (23,24) while APOC relied on the REMO (Rapid Epidemiological
260 Mapping of Onchocerciasis) (5,6,25,26) methodology to identify endemic areas requiring MDA, employing
- 260 Mapping of Onchocerciasis)(5,6,25,26) methodology to identify endemic areas requiring MDA, employing nodule palpation in adults. To harmonise these data sources, we adjusted survey data to represent all-age microfilade
- 261 palpation in adults. To harmonise these data sources, we adjusted survey data to represent all-age microfiladermia
- 262 prevalence by developing and applying an age and diagnostic crosswalk model that simultaneously accounts for
263 viological age-dependent trends in prevalence and differential detection sensitivity between skin snip ex
- 263 typical age-dependent trends in prevalence and differential detection sensitivity between skin snip examination and nodule palpation. We did not crosswalk results from antibody (eg. $Ov16$) surveys as these data were 264 nodule palpation. We did not crosswalk results from antibody (eg, Ov16) surveys as these data were not used in fitting the MBG model.
- fitting the MBG model.

 $\frac{266}{267}$ 267 We identified published within-study comparisons reporting results from skin snip biopsies (microfiladermia),
268 nodule (onchocercoma) palpation, or both, for more than one age group in a given study population. Study 268 nodule (onchocercoma) palpation, or both, for more than one age group in a given study population. Study cohorts with zero reported cases were excluded from the crosswalk training dataset because they do not contribute 269 with zero reported cases were excluded from the crosswalk training dataset because they do not contribute
270 information with which to derive age trends in endemic settings. Studies were eligible for inclusion in the 270 information with which to derive age trends in endemic settings. Studies were eligible for inclusion in the crosswalk 271 training dataset if they reported data from a country in our MBG modelling region; all eligible studies reported data
272 from surveys conducted in our geospatial modelling time frame (1988–2018). We identified 133 uni 272 from surveys conducted in our geospatial modelling time frame (1988–2018). We identified 133 unique survey
273 populations from 36 studies reporting skin snip prevalence in multiple age groups, and 126 unique survey 273 populations from 36 studies reporting skin snip prevalence in multiple age groups, and 126 unique survey
274 populations from 22 studies reporting nodule prevalence in multiple age groups; among these surveys, a to 274 populations from 22 studies reporting nodule prevalence in multiple age groups; among these surveys, a total of 100
275 unique survey populations from 19 studies reported both skin snip and nodule prevalence from the s 275 unique survey populations from 19 studies reported both skin snip and nodule prevalence from the same study
276 populations and age groups. Supplementary Table 6 summarises these sources. populations and age groups. Supplementary Table 6 summarises these sources.

277 **Supplementary Table 6. Data used in estimation of age and diagnostic crosswalk.**

278 Note that studies may have reported data for additional diagnostic tests that were not used in crosswalk model
279 fitting. fitting.

 $\frac{280}{281}$

281 We first retrieved population estimates by single age-year from the Global Burden of Disease (GBD)(66) for the country and year of each survey in the crosswalk training set. We assumed that the age distribution $(P(A))$, 282 country and year of each survey in the crosswalk training set. We assumed that the age distribution ($P(A)$, or probability of age A) within a survey sample matched that in the country and year of the survey, and estim 283 probability of age A) within a survey sample matched that in the country and year of the survey, and estimated prevalence-by-age models $(P(D|A))$, or the probability of disease, D, at age A) from birth through age 94 ye prevalence-by-age models ($P(D|A)$, or the probability of disease, D, at age A) from birth through age 94 years, the 285 maximum individual age-year modelled by GBD. We used the GBD age distributions to estimate the likelihood of observing the reported number of cases in each surveyed age bin, given a logistic (binomial) regression model 286 observing the reported number of cases in each surveyed age bin, given a logistic (binomial) regression model of the average prevalence-by-age relationship across surveys: average prevalence-by-age relationship across surveys:

288

 $\text{logit}(P(D|A)) = \beta_0 + I_{\text{ss}}f_{\text{ss}}(A) + I_{\text{nod}}f_{\text{nod}}(A) + \alpha_i + I_{\text{ss}}\beta_{\text{ss}}$

290
291 291 Onchocerciasis prevalence at a given age $(P(D|A))$ was modelled in logit space as a linear combination of an intercept, β_0 , which was set at logit(0.00001) to drive prevalence at birth toward zero; basis splines (fd

292 intercept, β_0 , which was set at logit(0.00001) to drive prevalence at birth toward zero; basis splines (fda R package(67)) on age, $f(A)$, to accommodate non-linear age trends; α_i , cohort-level fixed effects (in

package(67)) on age, $f(A)$, to accommodate non-linear age trends; α_i , cohort-level fixed effects (indicator variables

- 294 identifying study cohorts in the training dataset) to account for differences in study populations and survey designs;
295 and a fixed diagnostic effect for skin snip surveys, B_{sc} , to model overall differences b
- 295 and a fixed diagnostic effect for skin snip surveys, β_{ss} , to model overall differences between skin snip and nodule surveys. Diagnostic indicator variables, I_{ss} and I_{nod} , were set to 1 for skin snip and nodul
- 296 surveys. Diagnostic indicator variables, I_{ss} and I_{nod} , were set to 1 for skin snip and nodule observations, respectively, and set to 0 otherwise, to select the appropriate age and diagnostic effects for a given s

297 respectively, and set to 0 otherwise, to select the appropriate age and diagnostic effects for a given survey. Study
298 cohort effects were assumed to be shared across diagnostic tests. The likelihood function was eva

298 cohort effects were assumed to be shared across diagnostic tests. The likelihood function was evaluated as the sum,
299 coross studies (*i*) and age bins ($A1 \le A < A2$), of the binomial likelihood of observing the repor

299 across studies (*i*) and age bins ($A1 \le A < A2$), of the binomial likelihood of observing the reported numbers of cases given the reported sample sizes, age ranges, and diagnostic tests, using population-weighted mean p 300 cases given the reported sample sizes, age ranges, and diagnostic tests, using population-weighted mean prevalence valuated at the midpoints of each age year. values evaluated at the midpoints of each age year.

302
303

303 Spline knot placements were identified by spacing four internal knots evenly by quantile in the training data,
304 determined separately for skin snip and nodule surveys, with additional knots placed at ages 3, 6, and 304 determined separately for skin snip and nodule surveys, with additional knots placed at ages 3, 6, and 65 to help
305 stabilise model behaviour in early childhood and older adulthood. Starting values for all spline, co 305 stabilise model behaviour in early childhood and older adulthood. Starting values for all spline, cohort and
306 diagnostic coefficients were randomly drawn from uniform distributions in the interval [-5, 5]. Models we 306 diagnostic coefficients were randomly drawn from uniform distributions in the interval $[-5, 5]$. Models were estimated with maximum likelihood optimisation using the box constraint quasi-Newton algorithm⁶⁶ (withou 307 estimated with maximum likelihood optimisation using the box constraint quasi-Newton algorithm⁶⁶ (without constraints) in the optim function (R stats package(69)). constraints) in the optim function (R stats package(69)).

309
310

310 Prior to running the MBG model, all survey data that derived from a restricted age range (ie, anything other than 0–
311 94 years), or that represented nodule prevalence, were adjusted using the fitted crosswalk model. 311 94 years), or that represented nodule prevalence, were adjusted using the fitted crosswalk model. For each individual survey population, the crosswalk model was refit via maximum likelihood to the prevalence data for t 312 survey population, the crosswalk model was refit via maximum likelihood to the prevalence data for that survey, with all coefficients other than the cohort-level effect fixed, α_i , fixed to their estimates from the

313 with all coefficients other than the cohort-level effect fixed, α_i , fixed to their estimates from the full crosswalk model. This newly estimated value of α_i was then re-inserted into the crosswalk model with the

314 model. This newly estimated value of α_i was then re-inserted into the crosswalk model with the other coefficients 315 fixed, to calculate all-age skin snip prevalence for that survey sample. Crosswalked estimates

315 fixed, to calculate all-age skin snip prevalence for that survey sample. Crosswalked estimates were then used as the

outcome measure in the MBG model.

317
318 318 Surveys that reported prevalence of 0% or 100% were particularly difficult to interpret for crosswalk purposes. A report of 0% prevalence could reflect true absence of infection, insufficient diagnostic sensitivity, or 319 report of 0% prevalence could reflect true absence of infection, insufficient diagnostic sensitivity, or sampling
320 variance in small samples, while a report of 100% prevalence masks any implicit linkage between prev 320 variance in small samples, while a report of 100% prevalence masks any implicit linkage between prevalence and
321 the scale of infection intensity. While many of these concerns also apply to surveys reporting intermed 321 the scale of infection intensity. While many of these concerns also apply to surveys reporting intermediate
322 erevalence, the particular uncertainty involved with surveys reporting 0% or 100% prevalence increases the 322 prevalence, the particular uncertainty involved with surveys reporting 0% or 100% prevalence increases the risk of inappropriate crosswalking, and for this reason we did not crosswalk such surveys. inappropriate crosswalking, and for this reason we did not crosswalk such surveys.

324
325 325 Uncertainty in the crosswalk models was estimated using bootstrap analysis, with 100 replicates generated by sampling, with replacement, an equal number of study cohorts as in the full crosswalk training dataset. Resam 326 sampling, with replacement, an equal number of study cohorts as in the full crosswalk training dataset. Resampling was conducted by cohort rather than age bin to better estimate variation among surveys. Spline knot loc

327 was conducted by cohort rather than age bin to better estimate variation among surveys. Spline knot locations and
328 model coefficients were estimated for each bootstrap replicate independently of other replicates and

328 model coefficients were estimated for each bootstrap replicate independently of other replicates and of the full
329 model. To visualise bootstrapped results for a given survey, cohort effects were calculated for each

329 model. To visualise bootstrapped results for a given survey, cohort effects were calculated for each replicate using
330 maximum likelihood optimisation with all other coefficients fixed to the estimated values for the maximum likelihood optimisation with all other coefficients fixed to the estimated values for the replicate.

331

334 Outputs from the diagnostic and age crosswalk model. (a) Example prevalence-by-age curves fit to a study from
335 Benin(27) that reported both skin snip and nodule age-binned data and was included in the crosswalk mode 335 Benin(27) that reported both skin snip and nodule age-binned data and was included in the crosswalk model training
336 dataset. Estimates are shown for skin snip microfiladermia prevalence (solid black line) and nodule

336 dataset. Estimates are shown for skin snip microfiladermia prevalence (solid black line) and nodule prevalence (dotted green line); horizontal bars indicate the age-binned prevalence data reported by that study (solid

337 (dotted green line); horizontal bars indicate the age-binned prevalence data reported by that study (solid gray: skin 338 snip; dotted green: nodule). (b) Bootstraped estimates of differences between microfiladermia (s

338 snip; dotted green: nodule). **(b)** Bootstraped estimates of differences between microfiladermia (skin snip) and nodule

339 prevalence by age, in logit space. Each of 100 bootstrap samples is shown (faint grey line) along with the median estimate (solid black line) and central 95% uncertainty interval (shaded area). Positive values indicate 340 estimate (solid black line) and central 95% uncertainty interval (shaded area). Positive values indicate higher skin snip prevalence than nodule prevalence.

snip prevalence than nodule prevalence.

343 In an example location with moderately high microfiladermia prevalence, our final crosswalk model (Supplementary 344 Figure 5) suggests a sharp increase in prevalence through childhood and adolescence, with a slower in

344 Figure 5) suggests a sharp increase in prevalence through childhood and adolescence, with a slower increase
345 between approximately ages 15 and 65 years. This general pattern is qualitatively consistent with the rang

345 between approximately ages 15 and 65 years. This general pattern is qualitatively consistent with the range of onchocerciasis microfiladermia age trends reported by other studies, (70–72) although substantial site-spec 346 onchocerciasis microfiladermia age trends reported by other studies, $(70-72)$ although substantial site-specific
347 variation exists. The relationship between microfiladermia and nodule prevalence within onchocercias

347 variation exists. The relationship between microfiladermia and nodule prevalence within onchocerciasis-endemic
348 communities was previously modelled by Coffeng et al.,(73) using pre-control data from a broad sampling

348 communities was previously modelled by Coffeng et al.,(73) using pre-control data from a broad sampling of sites in
349 CCP and APOC regions. Their multivariate logistic regression model, relating nodule prevalence in

349 OCP and APOC regions. Their multivariate logistic regression model, relating nodule prevalence in adult males
350 (aged 20 years or older) to microfiladermia prevalence in individuals aged 5 years or older, did not exp 350 (aged 20 years or older) to microfiladermia prevalence in individuals aged 5 years or older, did not explicitly model

351 changes in prevalence with age, but revealed higher prevalence estimates from skin snip biopsies than those from
352 nodule palpation surveys in the same communities. Our full crosswalk model and bootstrapped uncertain

352 nodule palpation surveys in the same communities. Our full crosswalk model and bootstrapped uncertainty estimates
353 (Supplementary Figure 5) are qualitatively consistent with their results, estimating that skin snip

353 (Supplementary Figure 5) are qualitatively consistent with their results, estimating that skin snip prevalence exceeds nodule prevalence, on average, from adolescence until about age 50, with poorer resolution of this 354 nodule prevalence, on average, from adolescence until about age 50, with poorer resolution of this relationship in

- children and older adults.
- 356

357 While crosswalking survey data that are based on differing diagnostics and age coverage enabled us to leverage a
358 more comprehensive geospatial dataset than is otherwise tractable, we identify several limitations to 358 more comprehensive geospatial dataset than is otherwise tractable, we identify several limitations to our crosswalk
359 approach. Our model assumes that, for a given diagnostic approach, changes in prevalence by age fo 359 approach. Our model assumes that, for a given diagnostic approach, changes in prevalence by age follow a 360 consistent pattern across locations, years, sexes, and programmatic contexts. However, actual and reported microfiladermia and nodule prevalence patterns are influenced by local factors including ecological conditions; 362 infection intensity; vector identity, density, exposure, and control history; MDA coverage; variations in survey
363 sampling designs and in diagnostic specificity and sensitivity; and sex.(70,72–79) We have not modell 363 sampling designs and in diagnostic specificity and sensitivity; and sex.(70,72–79) We have not modelled the age-
364 and diagnostic-specific effects of MDA on prevalence in the crosswalk models (although MDA is include 364 and diagnostic-specific effects of MDA on prevalence in the crosswalk models (although MDA is included as a predictor during the MBG modelling stage). We were also limited by the variable reporting of age information 365 predictor during the MBG modelling stage). We were also limited by the variable reporting of age information among data sources. Some surveys did not report age ranges, and for these surveys we assumed that their data 366 among data sources. Some surveys did not report age ranges, and for these surveys we assumed that their data
367 represented all-age prevalence, risking possible misclassification. The absence of individual-level data 367 represented all-age prevalence, risking possible misclassification. The absence of individual-level data on
368 onchocerciasis prevalence also precluded full age-standardisation, as we could only assume that the survey 368 onchocerciasis prevalence also precluded full age-standardisation, as we could only assume that the survey sample
369 matched the age structure of the general population. Our crosswalks do not currently account for the 369 matched the age structure of the general population. Our crosswalks do not currently account for the sensitivity and 370 specificity of skin snip and nodule diagnostics, and crosswalk uncertainty is therefore likely 370 specificity of skin snip and nodule diagnostics, and crosswalk uncertainty is therefore likely to be underestimated.
371 We also do not currently have a computationally feasible method to propagate uncertainty from the 371 We also do not currently have a computationally feasible method to propagate uncertainty from the crosswalk
372 models into the MBG models in a way that accounts for the inconsistent reporting of sampled age ranges amo 372 models into the MBG models in a way that accounts for the inconsistent reporting of sampled age ranges among studies. studies. 374

375

376 **5.2 Polygon resampling** 377 Prevalence records are representative of either georeferenced point locations or polygonal areas (eg, as defined by
378 the borders of administrative or programmatic units). As our modelling framework relies on coordin 378 the borders of administrative or programmatic units). As our modelling framework relies on coordinate-referenced
379 data in order to fit the continuous spatial random fields, we converted areal data to a representativ 379 data in order to fit the continuous spatial random fields, we converted areal data to a representative collection of point data. This "polygon resampling" process, described previously for geostpatial modelling of unde 380 point data. This "polygon resampling" process, described previously for geostpatial modelling of under-5
381 mortality, (80) generates candidate locations based on the underlying population density of the resampled 381 mortality,(80) generates candidate locations based on the underlying population density of the resampled area,
382 implicitly assuming that surveys employed population-based designs, and is illustrated in Supplementary 382 implicitly assuming that surveys employed population-based designs, and is illustrated in Supplementary Figure 6
383 using Cameroon in 2015 as an example. For each polygonal observation in our dataset, 10 000 points we 383 using Cameroon in 2015 as an example. For each polygonal observation in our dataset, 10 000 points were randomly
384 sampled from within the polygon, with weighting by the WorldPop total population raster. Candidate po 384 sampled from within the polygon, with weighting by the WorldPop total population raster. Candidate points were
385 clustered using k-means clustering, generating a set of final points with a density of 1 per 1000 grid 385 clustered using k-means clustering, generating a set of final points with a density of 1 per 1000 grid cells, except for 386 small polygons, in which case density was iteratively increased by a factor of 10 until a min 386 small polygons, in which case density was iteratively increased by a factor of 10 until a minimum threshold of 10
387 points was achieved. Weights were assigned to each point proportionally to the number of candidate p 387 points was achieved. Weights were assigned to each point proportionally to the number of candidate points that 388 entered into the k-means cluster. The points generated by this resampling process were assigned the pre 388 entered into the k-means cluster. The points generated by this resampling process were assigned the prevalence of onchocerciasis reported for the survey for that polygon. These sample weights were used in MBG model fit 389 onchocerciasis reported for the survey for that polygon. These sample weights were used in MBG model fitting

390 within INLA.

Supplementary Figure 6. Polygon resampling.

394 The process of polygon resampling is illustrated using reported onchocerciasis prevalence data for Cameroon in 2015. (a) National map of Cameroon, showing the inset area (red frame) that is featured in the remaining pa 2015. (**a**) National map of Cameroon, showing the inset area (red frame) that is featured in the remaining panels. (**b**) Reported areal (irregular polygons) and point-level (circles) prevalence data. (**c**) Underlying population surface from 397 WorldPop (displayed on a log₁₀ scale), with survey polygons overlayed. Polygons are resampled proportionally to 5
398 x 5 km population density. (d) Locations of the final datapoints for geospatial modelling, showin x 5 km population density. (**d**) Locations of the final datapoints for geospatial modelling, showing both those data 399 originally reported with coordinates and those derived by resampling the polygon data. The opacity of resampled datapoints in (d) reflects their relative weights, which sum to 1.0 within an individual areal survey. datapoints in (d) reflects their relative weights, which sum to 1.0 within an individual areal survey.

401 **5.3 Geostatistical model**

402 **5.3.1 Model geographies and time period**

403 Model-based geostatistical (MBG) methods were used to generate estimates of all-age onchocerciasis microfiladermia prevalence for onchocerciasis-endemic countries of Africa, plus Yemen (listed in Sur

404 microfiladermia prevalence for onchocerciasis-endemic countries of Africa, plus Yemen (listed in Supplementary
405 Table 2). A single model was fit to this geographical region. We did not model countries in the Western

405 Table 2). A single model was fit to this geographical region. We did not model countries in the Western Hemisphere

- that were formerly or are currently endemic for onchocerciasis, due to their highly localised onchocerciasis foci. We 407 were principally concerned with estimates for the time period 2000–2018, but fit the model using d
- 407 were principally concerned with estimates for the time period 2000–2018, but fit the model using data from 1988–
408 2018 in order to leverage information from pre-2000 OCP and APOC surveys, and thereby improve "baseli
- 2018 in order to leverage information from pre-2000 OCP and APOC surveys, and thereby improve "baseline" (ie,
409 vear 2000) model estimates in countries covered by those programmes. Reporting of results in the main text f
- year 2000) model estimates in countries covered by those programmes. Reporting of results in the main text focuses
- on estimates for 2000–2018. $^{410}_{411}$

412 **5.3.2 Covariate coverage**

413 As with any regression model, the reliability of predictions from our model is affected by the overlap between
414 covariate values in training and prediction datasets. Predictions in regions with a range of covariate

- 414 covariate values in training and prediction datasets. Predictions in regions with a range of covariate values that fall outside the range of values in the training set may be prone to extrapolation errors. Supplementar
- 415 outside the range of values in the training set may be prone to extrapolation errors. Supplementary Figure 7
416 illustrates the number of covariates whose mean values (averaged over 1988–2018) fall outside the central
- 416 illustrates the number of covariates whose mean values (averaged over 1988–2018) fall outside the central 95% quantile interval of values in the training set, for each 5 x 5-km pixel. Child model predictions in areas w
- quantile interval of values in the training set, for each 5 x 5-km pixel. Child model predictions in areas with poor
- 418 covariate representation (ie, a large number of covariates with values outside the central interval) should be
419 considered with special caution. These areas include the Sahel and Sahara, as well as Yemen. Kenva, Som
- 419 considered with special caution. These areas include the Sahel and Sahara, as well as Yemen, Kenya, Somalia, 420 eastern Ethiopia, and southern Angola.
- eastern Ethiopia, and southern Angola. 421

422 **5.3.3 Environmental suitability**

- 423 Cromwell et al. recently produced an environmental suitability model for onchocerciasis, using a boosted regression tree (BRT) model trained with space- and time-referenced data on onchocerciasis occurrence. (81) In co
- 424 tree (BRT) model trained with space- and time-referenced data on onchocerciasis occurrence. (81) In contrast to our
425 present MBG model, the Cromwell suitability model leveraged data from not only skin snip and nodul
- 425 present MBG model, the Cromwell suitability model leveraged data from not only skin snip and nodule palpation
426 surveys, but also Ov16 serosurveys, onchocerciasis-derived eye or skin morbidity surveys, and other diag
- 426 surveys, but also Ov16 serosurveys, onchocerciasis-derived eye or skin morbidity surveys, and other diagnostics.
427 The model was trained using an overlapping but non-identical set of environmental covariates as those
- 427 The model was trained using an overlapping but non-identical set of environmental covariates as those used in the 428 present study, including climatic, topographic, hydrologic, vegetative, and urbanicity variables. Ps
- 428 present study, including climatic, topographic, hydrologic, vegetative, and urbanicity variables. Pseudo-absence
429 ecords were generated using background sampling. Importantly for our present purposes, this suitabili 429 records were generated using background sampling. Importantly for our present purposes, this suitability model does
430 not incorporate the influence of programmatic interventions for onchocerciasis and does not use da
- 430 not incorporate the influence of programmatic interventions for onchocerciasis and does not use data on onchocerciasis absence, providing a reflection of *O. volvulus* endemicity apart from temporal changes in
- 431 onchocerciasis absence, providing a reflection of *O. volvulus* endemicity apart from temporal changes in infection 432 prevalence. The outputs of the BRT model represent a spatially explicit index from 0–1, reflecting
- 432 prevalence. The outputs of the BRT model represent a spatially explicit index from $0-1$, reflecting environmental suitability for onchocerciasis occurrence in a given 5 x 5-km cell. We incorporated mean predictions f
- 433 suitability for onchocerciasis occurrence in a given 5×5 -km cell. We incorporated mean predictions from this suitability model, using covariate values for 2016 (Supplementary Figure 8), to improve the behaviour of
- 434 suitability model, using covariate values for 2016 (Supplementary Figure 8), to improve the behaviour of our MBG
435 model in areas of poor prevalence data coverage, particularly in areas where covariates lie outside t
- 435 model in areas of poor prevalence data coverage, particularly in areas where covariates lie outside the range of values in sampled locations and MBG predictions are therefore subject to extrapolation uncertainty. 436 values in sampled locations and MBG predictions are therefore subject to extrapolation uncertainty.

437
438
439 **Supplementary Figure 7. Number of covariates with values outside the central 95% interval of values at** survey sites.

 For each 5 x 5-km pixel, the number of covariates (out of a total of 16) whose average value lies outside the central 95% interval of values at observed survey sites is indicated. Localities with a large number of such covariates

represent areas in which predictions are less reliable, due to potential extrapolation error.

-
-

Supplementary Figure 8. Onchocerciasis suitability model predictions (Cromwell et al., 2021) for 2016.

449 Mean environmental suitability for onchocerciasis, from a previous boosted regression tree (BRT) analysis (81) of onchocerciasis presence data, using covariate values for 2016. 450 onchocerciasis presence data, using covariate values for 2016.
452 **5.3.4 Model description**

5.3.4 Model description

 We modelled *Onchocerca volvulus* infection prevalence using a spatially explicit Bayesian generalised linear mixed effects regression model in R-INLA:

455
456 $Y_{i,t} \sim \text{Binomial}(p_{i,t}, N_{i,t})$

457
458

$$
\logit(p_{i,t}) = \beta_0 + \beta X_{i,t} + \gamma_{c[i]} + U_i + Z_i + \epsilon_{i,t}
$$

- $Z_i \sim GP(0, \Sigma_{space})$
-

$$
\epsilon_{i,t} \sim N(0, \sigma_{nug}^2)
$$

- 463
464
- The number of infected individuals $(Y_{i,t})$ among a sample $(N_{i,t})$ in location *i* and year *t* was modelled as a binomial variable. This model specifies logit-transformed infection prevalence (p_i, t) as a linear combinati
- 465 variable. This model specifies logit-transformed infection prevalence $(p_{i,t})$ as a linear combination of an intercept 466 for the modelling region (β_0) ; covariate fixed effects (coefficients, β , and values, X_{i
- for the modelling region (β_0); covariate fixed effects (coefficients, β , and values, $X_{i,t}$); country random effects ($\gamma_{c[i]}}$); a second-order random walk model on estimates of onchocerciasis environmental suita
- ($\gamma_{c[i]}$); a second-order random walk model on estimates of onchocerciasis environmental suitability (U_i); a spatially
468 correlated random field (Z_i); and an uncorrelated error term or nugget effect (ϵ_{i+1}). Th
- 468 correlated random field (Z_i) ; and an uncorrelated error term or nugget effect $(\epsilon_{i,t})$. The spatial random field (Z_i) was modelled as a Gaussian process with mean 0 and a Matérn covariance function (Σ_{space}) . Modelled
- modelled as a Gaussian process with mean 0 and a Matérn covariance function (Σ_{space}). Modelled outputs from the 470 BRT analysis of onchocerciasis environmental suitability (U_i), described in section 5.3.3, were modell
- ATO BRT analysis of onchocerciasis environmental suitability (U_i) , described in section 5.3.3, were modelled with a second-order random walk, with values grouped into 25 bins by quantile. This random walk model accommoda
- 471 second-order random walk, with values grouped into 25 bins by quantile. This random walk model accommodates non-linearity in the relationship between suitability (based on presence-absence data) and the linear predicto
- 472 non-linearity in the relationship between suitability (based on presence-absence data) and the linear predictor. The INLA (integrated nested Laplace approximation) model was fit using an "empirical Bayes" integration s
- 473 INLA (integrated nested Laplace approximation) model was fit using an "empirical Bayes" integration strategy, 474 which relies on mode values to approximate hyperparameter posterior distributions during estimation and enabled us 475 to achieve tractable model computation times. (82)
- to achieve tractable model computation times.(82)
- 476

477 **5.3.5 Priors**

- 478 We specified minimally informative priors for INLA hyperparameters, as detailed in Supplementary Table 7. Priors 479 for the spatial hyperparameters τ and κ were derived automatically by R-INLA based on the f
- for the spatial hyperparameters τ and κ were derived automatically by R-INLA based on the finite elements mesh.

480 **Supplementary Table 7. INLA model priors.**

¹PC prior for precision.

²PC prior for SPDE σ, indicating a 5% probability that σ is greater than 3.

³PC prior for SPDE range, indicating a 5% probability that range is less than 0.06161412 (5% of the maximum extent of the spatial mesh).

481

482 **5.3.6 Mesh construction**

483 We modelled continuous spatial random effects using stochastic partial differential equations (SPDE)
484 representations of Gaussian-Markov random field (GMRF) approximations of a spatially autocorrelate

484 representations of Gaussian-Markov random field (GMRF) approximations of a spatially autocorrelated Gaussian
485 process, using triangular finite element meshes as implemented in the R-INLA R package. (83–85) Due to th

485 process, using triangular finite element meshes as implemented in the R-INLA R package.(83–85) Due to the large 486 geographical size of the model region, a spherical (S2) mesh was constructed in order to minimise dist

486 geographical size of the model region, a spherical (S2) mesh was constructed in order to minimise distance distance distortions. Minimum and maximum edge lengths were set to 25 and 500 km, respectively, and a 1000-km

487 distortions. Minimum and maximum edge lengths were set to 25 and 500 km, respectively, and a 1000-km external
488 buffer was used to avoid artifacts at the edges of the spatial domain. (84) These values were chosen to

488 buffer was used to avoid artifacts at the edges of the spatial domain.(84) These values were chosen to yield denser
489 mesh vertices in data-rich areas while maximising the overall spatial field resolution, within com

489 mesh vertices in data-rich areas while maximising the overall spatial field resolution, within computational constraints. The spatial mesh is illustrated in Supplementary Figure 9.

490 constraints. The spatial mesh is illustrated in Supplementary Figure 9.

491

492
493

493 **Supplementary Figure 9. Spatial mesh construction.**

494 Two-dimensional projection of spherical refined Delaunay triangulation mesh used in estimating spatial random
495 fields in Africa and Yemen, with national boundaries (bold lines). The mesh features greater vertex dens 495 fields in Africa and Yemen, with national boundaries (bold lines). The mesh features greater vertex density in data-
496 rich locations. rich locations.

497

498 **5.3.7 Model fitting and estimation generation** 499 Models were fit using the integrated nested Laplace approximation (INLA) algorithm in R-INLA. Fitted models for
500 each region were used to generate 1000 random samples from the joint posterior distributions of model 500 each region were used to generate 1000 random samples from the joint posterior distributions of model parameters,
501 vielding mean and uncertainty estimates for onchoerciasis prevalence.

- yielding mean and uncertainty estimates for onchocerciasis prevalence.
- 502

503 **5.3.8 Model results**

- 504 Model parameter estimates from the MBG model are summarised in Supplementary Table 8. Nominal range is the
505 distance (in km) at which spatial correlation has declined to about 0.1. Estimated random effects for the
- 505 distance (in km) at which spatial correlation has declined to about 0.1. Estimated random effects for the environmental suitability covariate are displayed in Supplementary Figure 10.
- 506 environmental suitability covariate are displayed in Supplementary Figure 10.

507 **Supplementary Table 8. Parameter estimates from in-sample onchocerciasis MBG model.**

514 **Supplementary Figure 10. Estimated random effects for environmental suitability.**

515 Random effect estimates (mean and 95% UI) are shown for the environmental suitability layer, fit within INLA
516 using a second-order random walk (RW2) model with standard-normalised suitability values (x-axis) grouped 516 using a second-order random walk (RW2) model with standard-normalised suitability values (x-axis) grouped into 517 25 bins by quantile; y-axis values indicate effects in logit space. 517 25 bins by quantile; y-axis values indicate effects in logit space.

518

519 **5.3.9** *Loa loa* **endemicity** 520 The co-occurrence of *O. volvulus* and the filarial nematode *Loa loa* complicates onchocerciasis control in some parts of central Africa,(86,87) due to potentially severe complications from ivermectin treatment in ind 521 of central Africa,(86,87) due to potentially severe complications from ivermectin treatment in individuals with high 522 L. loa microfilariae loads.(88,89) Previous modelling studies have estimated the scale of co-i 522 *L. loa* microfilariae loads.(88,89) Previous modelling studies have estimated the scale of co-infections between *O.* 523 *volvulus* and *L. loa* and have suggested a substantial burden of co-infection or populations at risk.(90,91) We 524 complemented these studies by calculating the mean number of onchocerciasis cases estimated by our geospatial
525 model in 2018, in areas considered meso- or hyper-endemic for loiasis according to endemicity classifica 525 model in 2018, in areas considered meso- or hyper-endemic for loiasis according to endemicity classifications from
526 ESPEN. We obtained data from ESPEN for loiasis endemicity in 2015 at the level of Implementation Un 526 ESPEN. We obtained data from ESPEN for loiasis endemicity in 2015 at the level of Implementation Units and
527 combined this with the WorldPop population raster and our onchocerciasis mean prevalence estimates, derivin 527 combined this with the WorldPop population raster and our onchocerciasis mean prevalence estimates, deriving a
528 mean estimate of 7 146 618 onchocerciasis cases in loiasis-endemic areas (excluding hypo-endemic region mean estimate of 7 146 618 onchocerciasis cases in loiasis-endemic areas (excluding hypo-endemic regions) in 529 2018.

530
531 531 **5.4 Model validation**

532 **5.4.1 Metrics of predictive validity**

- 533 In order to assess the predictive validity of our estimates, we validated our models using spatially stratified five-fold
534 out-of-sample cross-validation. To construct each spatial fold, we used a modified bi-tree a
- 534 out-of-sample cross-validation. To construct each spatial fold, we used a modified bi-tree algorithm to spatially
- 535 aggregate datapoints. This algorithm recursively partitions two-dimensional space, alternating between horizontal
- 536 and vertical splits on the weighted data sample size medians, until the data contained within each spatial partition are of a similar sample size. The depth of recursive partitioning is constrained by the target sample
- 537 of a similar sample size. The depth of recursive partitioning is constrained by the target sample size within a
538 partition and the minimum number of clusters or pseudo-clusters allowed within each spatial partition
- 538 partition and the minimum number of clusters or pseudo-clusters allowed within each spatial partition (in this case, a minimum sample size of 500 was used). These spatial partitions are then allocated to one of five fo
- 539 minimum sample size of 500 was used). These spatial partitions are then allocated to one of five folds for cross-
540 validation. Temporal partitioning was unstructured (random). validation. Temporal partitioning was unstructured (random).
-
- 541
542
- 542 For validation, each geostatistical model was run five times, each time holding out data from one of the folds. A set of out-of-sample predictions were generated by sampling from the posterior predictive distribution f
- 543 of out-of-sample predictions were generated by sampling from the posterior predictive distribution for each held-out datapoint. A full suite of out-of-sample predictions over the entire dataset was calculated by combin 544 datapoint. A full suite of out-of-sample predictions over the entire dataset was calculated by combining the out-of-
545 sample predictions from the five cross-validation runs. Using these out-of-sample predictions, we
- 545 sample predictions from the five cross-validation runs. Using these out-of-sample predictions, we computed mean
546 error (bias), mean absolute error, 95% coverage of the predictive intervals (the proportion of observe
- 546 error (bias), mean absolute error, 95% coverage of the predictive intervals (the proportion of observed out-of-sample
547 data that fall within the predicted 95% uncertainty intervals), root-mean squared error (RMSE, w
- 547 data that fall within the predicted 95% uncertainty intervals), root-mean squared error (RMSE, which summarises error variance), and the correlation of predicted versus observed prevalence at the level of individual da
- 548 error variance), and the correlation of predicted versus observed prevalence at the level of individual datapoints. A scatterplot of reported prevalence versus mean out-of-sample predictions is provided in Supplementar
- 549 scatterplot of reported prevalence versus mean out-of-sample predictions is provided in Supplementary Figure 11,
550 and validation metrics are summarised in Supplementary Table 9. and validation metrics are summarised in Supplementary Table 9.
-
- 551
552
- 552 In addition to performing cross-validation, we also evaluated over-dispersion in the model input data by performing
553 posterior predictive checks. Briefly, a binomial count was simulated from each of 1,000 model draw
- 553 posterior predictive checks. Briefly, a binomial count was simulated from each of 1,000 model draws of predictions 554 for each input data row. The distribution of predictions provided a good approximation of the obser 554 for each input data row. The distribution of predictions provided a good approximation of the observed distribution, suggesting that the model is adequately addressing possible over-dispersion in the data.
- suggesting that the model is adequately addressing possible over-dispersion in the data.

Supplementary Figure 11. Model validation scatterplots for Africa and Yemen.

559 Reported prevalence versus mean out-of-sample predictions for individual datapoints, by year and country. Vertical
560 bars represent 95% UI; red lines indicate equivalence. bars represent 95% UI; red lines indicate equivalence.

Supplementary Table 9. Out-of-sample validation metrics at the level of individual datapoints, from five-fold cross-validation

563 Out-of-sample performance was aggregated over 1988–2018 and is also provided for individual model years. Values were computed in prevalence space. N observations: Number of data rows in the full dataset from a given ye were computed in prevalence space. N observations: Number of data rows in the full dataset from a given year.

Mean abs. error: Mean absolute error. 95% cov.: 95% coverage. Corr.: Correlation. $\frac{565}{566}$

 3. Remme JHF. The African programme for onchocerciasis control: preparing to launch. Parasitol Today. 1995;11(11):403–6.

 4. Ngoumou P, Walsh JF, Mace JM. A rapid mapping technique for the prevalence and distribution of onchocerciasis: a Cameroon case study. Ann Trop Med Parasitol. 1994;88(5):463–74.

- 578 5. Ngoumou P, Walsh JF, Blindness WP for the P of, Diseases UBSP for R and T in T. A manual for rapid epidemiological mapping of onchocerciasis. 1993; Available from: https://apps.who.int/iris/handle/10665/59537
- 6. Noma M, Nwoke BEB, Nutall I, Tambala PA, Enyong P, Namsenmo A, et al. Rapid epidemiological mapping of onchocerciasis (REMO): its application by the African Programme for Onchocerciasis Control (APOC). Ann Trop Med Parasitol. 2002;96 Suppl 1:S29-39.

 7. Tekle AH, Zouré HGM, Noma M, Boussinesq M, Coffeng LE, Stolk WA, et al. Progress towards onchocerciasis elimination in the participating countries of the African Programme for Onchocerciasis Control: epidemiological evaluation results. Infect Dis Poverty. 2016;5(1):66.

 2. WHO | Geographical distribution of onchocerciasis. WHO. World Health Organization; Available from: https://www.who.int/onchocerciasis/distribution/en/

 8. WHO | African Programme for Onchocerciasis Control: progress report, 2013–2014. Wkly Epidemiol Rec. 2014;89(49):551–60.

- 9. ENVISION. Tanzania Work Plan, FY2018. ENVISION; 2017. Available from: https://www.ntdenvision.org/sites/default/files/docs/tanzania_work_planfy18.pdf
- 10. de Smet E, Metanmo S, Mbelesso P, Kemata B, Siewe Fodjo JN, Boumédiène F, et al. Focus of ongoing onchocerciasis transmission close to Bangui, Central African Republic. Pathogens. 2020;9(5):337.
- 11. Bof JCM, Maketa V, Bakajika DK, Ntumba F, Mpunga D, Murdoch ME, et al. Onchocerciasis control in the Democratic Republic of Congo (DRC): challenges in a post-war environment. Trop Med Int Health. 2015;20(1):48–62.
- 12. Makenga Bof JC, Ntumba Tshitoka F, Muteba D, Mansiangi P, Coppieters Y. Review of the National Program for Onchocerciasis Control in the Democratic Republic of the Congo. Trop Med Infect Dis. 2019;4(2):92.
- 13. Samuel A, Belay T, Yehalaw D, Taha M, Zemene E, Zeynudin A. Impact of six years community directed treatment with ivermectin in the control of onchocerciasis, Western Ethiopia. PLOS ONE. 2016;11(3):e0141029.
- 14. Gemade EII, Jiya JY, Nwoke BEB, Ogunba EO, Edeghere H, Akoh JI, et al. Human onchocerciasis: current assessment of the disease burden in Nigeria by rapid epidemiological mapping. Ann Trop Med Parasitol. 1998;92(sup1):S79–83.
- 15. Republic of Rwanda Ministry of Health. Neglected Tropical Diseases Strategic Plan 2019- 2024. Republic of Rwanda Ministry of Health; 2019. Available from: http://rbc.gov.rw/fileadmin/user_upload/guide2019/guide2019/RWANDA%20NTD%20ST RATEGIC%20PLAN%202019-2024.pdf
- 16. Laggo MCL, Chane F. Onchocerciasis control in South Sudan. South Sudan Med J. 2011;4(3):61–2.
- 17. World Bank. Chad, Cameroon Petroleum Development and Pipeline Project: environmental assessment. Washington, D.C.: World Bank; 1999. Available from: http://documents.banquemondiale.org/curated/fr/863021468769828252/pdf/multi-page.pdf
- 18. Katabarwa MN, Lakwo T, Habomugisha P, Unnasch TR, Garms R, Hudson-Davis L, et al. After 70 years of fighting an age-old scourge, onchocerciasis in Uganda, the end is in sight. 619 Int Health. 2018;10(suppl 1):i79–88.
- 19. ESPEN Portal. Available from: https://admin.espen.afro.who.int/
- 20. WHO | Global Programme to Eliminate Lymphatic Filariasis. WHO. World Health Organization; Available from: http://www.who.int/lymphatic_filariasis/elimination-programme/en/
- 21. Graham MH. Confronting Multicollinearity in Ecological Multiple Regression. Ecology. 2003;84(11):2809–15.
- 22. Faraway JJ. Linear Models with R. Boca Raton: CRC Press; 2005.
- 23. Prost A, Thylefors B, Pairault C. Methods of mass epidemiological evaluation of onchocerciasis: their utilization in a vector control programme. Expert Comm Epidemiol Onchocerciasis. 1975;20.
- 24. O'Hanlon SJ, Slater HC, Cheke RA, Boatin BA, Coffeng LE, Pion SDS, et al. Model-based geostatistical mapping of the prevalence of Onchocerca volvulus in West Africa. PLoS Negl Trop Dis. 2016;10(1):e0004328.
- 25. World Health Organization. Supplemental Guidelines for Rapid Epidemiological Mapping of Onchocerciasis (REMO). Geneva, Switzerland: World Health Organization; 1995. Report No.: Document TDR/TDF/ONCHO/95.1. Available from: https://www.who.int/tdr/publications/tdr-research-publications/oncho-guidelines/en/
-
- 26. Zouré HG, Noma M, Tekle AH, Amazigo UV, Diggle PJ, Giorgi E, et al. The geographic distribution of onchocerciasis in the 20 participating countries of the African Programme for Onchocerciasis Control: (2) pre-control endemicity levels and estimated number infected. Parasit Vectors. 2014;7(1):326.
- 27. Gallin M, Adams A, Kruppa TF, Gbaguidi EA, Massougbodji A, Sadeler BC, et al. Epidemiological studies of onchocerciasis in southern Benin. Trop Med Parasitol Off Organ Dtsch Tropenmedizinische Ges Dtsch Ges Tech Zusammenarbeit GTZ. 1993;44(2):69–74.
- 28. Boussinesq M, Chippaux JP, Ernould JC, Quillevere D, Prod'hon J. Effect of repeated treatments with ivermectin on the incidence of onchocerciasis in northern Cameroon. Am J Trop Med Hyg. 1995;53(1):63–7.
- 29. Mendoza Aldana J, Piechulek H, Maguire J. Forest onchocerciasis in Cameroon: its distribution and implications for selection of communities for control programmes. Ann Trop Med Parasitol. 1997;91(1):79–86.
- 30. Katabarwa MN, Eyamba A, Nwane P, Enyong P, Kamgno J, Kueté T, et al. Fifteen years of annual mass treatment of onchocerciasis with ivermectin have not interrupted transmission in the west region of Cameroon. J Parasitol Res. 2013;2013. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3652197/
- 31. Kamgno J, Boussinesq M. Hyperendemic loaiasis in the Tikar plain, shrub savanna region of Cameroon. Bull Soc Pathol Exot 1990. 2001;94(4):342–6.
- 32. Katabarwa MN, Eyamba A, Chouaibou M, Enyong P, Kuété T, Yaya S, et al. Does onchocerciasis transmission take place in hypoendemic areas? a study from the North Region of Cameroon. Trop Med Int Health. 2010;15(5):645–52.
- 33. Katabarwa MN, Eyamba A, Nwane P, Enyong P, Yaya S, Baldiagaï J, et al. Seventeen years of annual distribution of ivermectin has not interrupted onchocerciasis transmission in North Region, Cameroon. Am J Trop Med Hyg. 2011;85(6):1041–9.

 34. Kamga HLF, Shey DN, Assob JCN, Njunda AL, Nde Fon P, Njem PK. Prevalence of onchocerciasis in the Fundong Health District, Cameroon after 6 years of continuous community-directed treatment with ivermectin. Pan Afr Med J. 2011;10. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3240934/ 35. Wanji S, Kengne-Ouafo JA, Esum ME, Chounna PWN, Adzemye BF, Eyong JEE, et al. Relationship between oral declaration on adherence to ivermectin treatment and parasitological indicators of onchocerciasis in an area of persistent transmission despite a decade of mass drug administration in Cameroon. Parasit Vectors. 2015;8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4696282/ 36. Kamga GR, Dissak-Delon FN, Nana-Djeunga HC, Biholong BD, Ghogomu SM, Souopgui J, et al. Important progress towards elimination of onchocerciasis in the West Region of Cameroon. Parasit Vectors. 2017;10. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5543544/ 37. Kamga GR, Dissak-Delon FN, Nana-Djeunga HC, Biholong BD, Mbigha-Ghogomu S, Souopgui J, et al. Still mesoendemic onchocerciasis in two Cameroonian community- directed treatment with ivermectin projects despite more than 15 years of mass treatment. Parasit Vectors. 2016;9(1):581. 38. Bakajika D, Senyonjo L, Enyong P, Oye J, Biholong B, Elhassan E, et al. On-going transmission of human onchocerciasis in the Massangam health district in the West Region of Cameroon: Better understanding transmission dynamics to inform changes in programmatic interventions. PLoS Negl Trop Dis. 2018;12(11):e0006904. 39. Katabarwa M, Eyamba A, Habomugisha P, Lakwo T, Ekobo S, Kamgno J, et al. After a decade of annual dose mass ivermectin treatment in Cameroon and Uganda, onchocerciasis transmission continues. Trop Med Int Health TM IH. 2008;13(9):1196–203. 40. Kennedy MH, Bertocchi I, Hopkins AD, Meredith SE. The effect of 5 years of annual treatment with ivermectin (Mectizan) on the prevalence and morbidity of onchocerciasis in the village of Gami in the Central African Republic. Ann Trop Med Parasitol. 2002;96(3):297–307. 41. Taye A, Gebre-Michael T, Taticheff S. Onchocerciasis in Gilgel Ghibe River Valley southwest Ethiopia. East Afr Med J. 2000;77(2):116–20. 42. Mengistu G, Balcha F, Britton S. Co-infection of Onchocerca volvulus and intestinal helminths in indigenous and migrant farmers in southwest Ethiopia. Ethiop Med J. 2002;40(1):19–27. 43. Legesse M, Balcha F, Erko B. Status of onchocerciasis in Teppi area, Southwestern Ethiopia, after four years of annual community-directed treatment with ivermectin. Ethiop J Health Dev. 2010;24(1). Available from: https://www.ajol.info/index.php/ejhd/article/view/62945

 44. Dori GU, Belay T, Belete H, Panicker KN, Hailu A. Parasitological and clinico- epidemiological features of onchocerciasis in West Wellega, Ethiopia. J Parasit Dis Off Organ Indian Soc Parasitol. 2012;36(1):10–8. 45. Dana D, Debalke S, Mekonnen Z, Kassahun W, Suleman S, Getahun K, et al. A community-based cross-sectional study of the epidemiology of onchocerciasis in unmapped villages for community directed treatment with ivermectin in Jimma Zone, southwestern Ethiopia. BMC Public Health. 2015;15:595. 46. Fobi G, Mourou Mbina JR, Ozoh G, Kombila M, Agaya C, Olinga Olinga JM, et al. Onchocerciasis in the area of Lastourville, Gabon. Clinical and entomological aspects. Bull Soc Pathol Exot 1990. 2006;99(4):269–71. 47. Anosike JC, Onwuliri CO. Studies on filariasis in Bauchi State, Nigeria. II. The prevalence of human filariasis in Darazo Local Government area. Appl Parasitol. 1994;35(4):242–50. 48. Osue HO, Inabo HI, Yakubu SE, Audu PA, Galadima M, Odama LE, et al. Impact of eighteen-year varied compliance to onchocerciasis treatment with ivermectin in sentinel savannah agrarian communities in Kaduna State of Nigeria. ISRN Parasitol. 2013;2013:960168. 49. Anosike JC, Celestine null, Onwuliri OE, Onwuliri VA. The prevalence, intensity and clinical manifestations of Onchocerca volvulus infection in Toro local government area of Bauchi State, Nigeria. Int J Hyg Environ Health. 2001;203(5–6):459–64. 50. Umeh RE. The causes and profile of visual loss in an onchocerciasis-endemic forest- savanna zone in Nigeria. Ophthalmic Epidemiol. 1999;6(4):303–15. 51. Nmorsi OPG, Oladokun I a. A, Egwunyenga OA, Oseha E. Eye lesions and onchocerciasis in a rural farm settlement in Delta state, Nigeria. Southeast Asian J Trop Med Public 722 Health. 2002;33(1):28–32. 52. Okoye IC, Onwuliri CO. Epidemiology and psycho-social aspects of onchocercal skin diseases in northeastern Nigeria. Filaria J. 2007;6:15. 53. Rebecca SN, Akinboye DO, Abdulazeez AA. Onchocerciasis and plasmodiasis: concurrent infection in Garaha-Dutse community, Adamawa State Nigeria. Biomed Res. 2008;19(2). Available from: https://www.alliedacademies.org/abstract/onchocerciasis-and- plasmodiasis-concurrent-infectionrnin-garahadutse-community-adamawa-state-nigeria- 802.html 54. Emukah E, Nimzing J, Ezeabikwa M, Obiezu J, Okpala N, Miri E, et al. Impact assessment of repeated annual ivermectin on ocular and clinical onchocerciasis 14 years of annual mass drug administration in eight sentinel villages, southeast Nigeria. Am J Trop Med Hyg. 2010;83(Suppl 5):22. 55. Akinbo FO, Okaka CE. Hyperendemicity of onchocerciasis in Ovia Northeast Local Government Area, Edo State, Nigeria. East Afr J Public Health. 2010;7(1):84–6.

 67. Ramsay JO, Wickham H, Graves S, Hooker G. fda: Functional Data Analysis. 2018. Available from: https://CRAN.R-project.org/package=fda 68. Byrd RH, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. SIAM J Sci Comput. 1995;16(5):1190–208. 69. R: The R Project for Statistical Computing. Available from: https://www.r-project.org/ 70. Kirkwood B, Smith P, Marshall T, Prost A. Variations in the prevalence and intensity of microfilarial infections by age, sex, place and time in the area of the Onchocerciasis Control Programme. Trans R Soc Trop Med Hyg. 1983;77(6):857–61. 71. Vivas-Martínez S, Basáñez MG, Botto C, Rojas S, García M, Pacheco M, et al. Amazonian onchocerciasis: parasitological profiles by host-age, sex, and endemicity in southern Venezuela. Parasitology. 2000;121(5):513–25. 72. Anderson RM, Basáñez MG, Boussinesq M. Population biology of human onchocerciasis. Philos Trans R Soc Lond B Biol Sci. 1999;354(1384):809–26. 73. Coffeng LE, Pion SDS, O'Hanlon S, Cousens S, Abiose AO, Fischer PU, et al. Onchocerciasis: the pre-control association between prevalence of palpable nodules and skin microfilariae. PLoS Negl Trop Dis. 2013;7(4):e2168. 74. Hamley JID, Milton P, Walker M, Basáñez MG. Modelling exposure heterogeneity and density dependence in onchocerciasis using a novel individual-based transmission model, 792 EPIONCHO-IBM: Implications for elimination and data needs. PLoS Negl Trop Dis.
2019:13(12):e0007557. 2019;13(12):e0007557. 75. Gebrezgabiher G, Mekonnen Z, Yewhalaw D, Hailu A. Reaching the last mile: main challenges relating to and recommendations to accelerate onchocerciasis elimination in Africa. Infect Dis Poverty. 2019;8(1):60. 76. Vlaminck J, Fischer PU, Weil GJ. Diagnostic tools for onchocerciasis elimination programs. Trends Parasitol. 2015;31(11):571–82. 77. Turner HC, Walker M, Churcher TS, Basáñez MG. Modelling the impact of ivermectin on River Blindness and its burden of morbidity and mortality in African Savannah: EpiOncho projections. Parasit Vectors. 2014;7(1):241. 78. Remme. The global burden of onchocerciasis in 1990. Geneva, Switzerland: World Health Organization; 2004 p. 26. Available from: 804 https://www.who.int/apoc/publications/bod onchocerciasis1990 the global burden of on chocerciasis_in_1990.pdf 79. Unnasch TR, Golden A, Cama V, Cantey PT. Diagnostics for onchocerciasis in the era of elimination. Int Health. 2018;10(suppl_1):i20–6.

- 80. Golding N, Burstein R, Longbottom J, Browne AJ, Fullman N, Osgood-Zimmerman A, et al. Mapping under-5 and neonatal mortality in Africa, 2000–15: a baseline analysis for the Sustainable Development Goals. The Lancet. 2017;390(10108):2171–82.
- 81. Cromwell EA, Osborne JCP, Unnasch TR, Basáñez MG, Gass KM, Barbre KA, et al. Predicting the environmental suitability for onchocerciasis in Africa as an aid to elimination planning. PLoS Negl Trop Dis. 2021;15(7):e0008824.
- 814 82. Gómez-Rubio V, Palmí-Perales F. Multivariate posterior inference for spatial models with the integrated nested Laplace approximation. J R Stat Soc Ser C Appl Stat. 2019;68(1):199–215.
- 83. Lindgren F, Rue H, Lindström J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Stat Methodol 819 Ser B. 2011;73(4):423–98.
- 820 84. Blangiardo M, Cameletti M, Baio G, Rue H. Spatial and spatio-temporal models with R-821 INLA. Spat Spatio-Temporal Epidemiol. 2013;7:39–55.
- 85. Lindgren F, Rue H. Bayesian Spatial Modelling with R- INLA. J Stat Softw. 2015;63(19):1–25.
- 86. Wanji S, Ndongmo WPC, Fombad FF, Kengne-Ouafo JA, Njouendou AJ, Tchounkeu YFL, et al. Impact of repeated annual community directed treatment with ivermectin on loiasis parasitological indicators in Cameroon: Implications for onchocerciasis and lymphatic filariasis elimination in areas co-endemic with Loa loa in Africa. PLoS Negl Trop Dis. 828 2018;12(9):e0006750.
- 87. Zouré HGM, Wanji S, Noma M, Amazigo UV, Diggle PJ, Tekle AH, et al. The geographic distribution of Loa loa in Africa: results of large-scale implementation of the rapid assessment procedure for loiasis (RAPLOA). PLoS Negl Trop Dis. 2011;5(6):e1210.
- 88. Gardon J, Gardon-Wendel N, Demanga-Ngangue, Kamgno J, Chippaux JP, Boussinesq M. Serious reactions after mass treatment of onchocerciasis with ivermectin in an area endemic for Loa loa infection. The Lancet. 1997;350(9070):18–22.
- 89. Boussinesq M, Gardon J, Gardon-Wendel N, Chippaux JP. Clinical picture, epidemiology and outcome of Loa-associated serious adverse events related to mass ivermectin treatment of onchocerciasis in Cameroon. Filaria J. 2003;2(1):S4.
- 90. Vinkeles Melchers NVS, Coffeng LE, Boussinesq M, Pedrique B, Pion SDS, Tekle AH, et al. Projected number of people with onchocerciasis–loiasis coinfection in Africa, 1995 to 2025. Clin Infect Dis. 2020;70(11):2281–9.
- 91. Cano J, Basáñez MG, O'Hanlon SJ, Tekle AH, Wanji S, Zouré HG, et al. Identifying co- endemic areas for major filarial infections in sub-Saharan Africa: seeking synergies and preventing severe adverse events during mass drug administration campaigns. Parasit Vectors. 2018;11(1):70.