Stochastic models indicate rapid smallpox spread and mass mortality of Indigenous Australians after colonial exposure—Supplementary Material

Matthew C. Nitschke[1,2,3], Alan N. Williams[4,5,2], Shane D. Ingrey[6,7,2,3], Billy Griffiths[8], Nicholas Pitt[9], Lynette Russell[10,2,3], Sean Ulm[2], Kirsty Beller[7], Michael I. Bird[1,2,3], Syeda H. Fatima[1,12], Ian J. McNiven[10,2,3], Frédérik Saltré[1,2,3], Alison Bashford[9,2], Christopher Wilson[13,2,3], and Corey J. A. Bradshaw[1,2,3]

[1]Global Ecology | *Partuyarta Ngadluku Wardli Kuu*, College of Science and Engineering, Flinders University, Adelaide, South Australia, Australia;
[2]Centre of Excellence for Indigenous and Environmental Histories and Futures, James Cook University, Cairns, Queensland, Australia;
[3]ARC Centre of Excellence for Australian Biodiversity and Heritage, Wollongong, New South Wales, Australia;
[4]School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales, Australia;
[5]EMM Consulting, 20 Chandos Street, St Leonards, New South Wales, Australia;
[6]School of Earth, Atmospheric and Life Sciences, University of Wollongong, Wollongong, NSW, Australia;
[7]Gujaga Foundation, 1 Elaroo Avenue, La Perouse, New South Wales, Australia;
[8]Alfred Deakin Institute, Deakin University, Burwood, Victoria, Australia;
[9]Laureate Centre for History and Population, University of New South Wales, Sydney, New South Wales, Australia;
[10]Monash Indigenous Studies Centre, Monash University, Clayton, Victoria, Australia;
[11]College of Arts, Society and Education, James Cook University, Cairns, Queensland, Australia;
[12]The Kids Research Institute Australia, Nedlands, Western Australia, Australia;
[13]College of Arts, Law and Education, University of Tasmania, Hobart, Tasmania, Australia

A Parameter estimates

A.0.1 Smallpox characteristics

Most of what we know about smallpox comes from the meticulous observations and detailed laboratory studies done during the global campaign to eradicate the disease from 1967 to 1979. Three distinct variants of the disease were identified: ordinary, malignant, and haemorrhagic. The most common was ordinary-type smallpox, which occurred in 88.8% of unvaccinated subjects studied in a hospital in Madras (Chennai), India[1] (Table S1). The severity of smallpox in unvaccinated cases was typically affected by the victim's age and gender. Symptoms were more severe in pregnant women (often pregnancies would end even if the mother survived) and young children < 5 years old. The case-fatality rate for women was 40.8%, 61.1% of which was for pregnant women and 34.5% for non-pregnant women. In men, the fatality rate was lower at 30.2%. Because epidemics often disproportionately affected women, smallpox had an unusually long-lasting effect on population trends.[2]

Table S1: Smallpox types. Smallpox manifests in several different ways that vary in frequency and fatality rate. These numbers are based on a case study in 1972[1]

smallpox type	fatality (%)		frequency(%)	
	unvaccinated	vaccinated	unvaccinated	vaccinated
ordinary	9.3	0.7	42.1	58.4
ordinary confluent	62	26.3	22.8	4.6
ordinary semi-confluent	37	8.4	23.9	7
early haemorrhagic	100	100	0.7	1.4
late haemorrhagic	96.8	89.8	1.7	2.0

The infectious period of smallpox lasted 23 days on average, which included four distinct stages with different infectivity (Table S2). To account for the varying infectivity and contact between individuals during the course of the infectious period, we reduced the length of the infectious stage in our model. Instead of the full mean of 23 days, we set the infectious period to 18 days. This includes the full 14 days of the most infectious stage, one day of the least infectious stage, and 3 days of the intermediate infectious stage at recovery.

Table S2: Main infectious states of smallpox. The most infectious period lasts an average of 14 days, with less infectious stages on either side. Those in stages 2 and 3 are extremely ill and need constant care. All those who die from this disease do so during stage 3.[3]

Stage	Details	Average duration	Infectivity
1	initial symptoms (fever, headache)	2-4 days	low
2	early rash	4 days	high
3	pustular rash & scabs (death)	10 days	highest
4	scabs fall off/recovery	6 days	medium

Table S3: Parameters and estimates of the stochastic model.

Parameter	Description	Estimate
a	proportion of population moving for regular trade	variable
b	proportion of population moving for ceremony	variable
μ	probability of movement outside home patch	Eq.(1)
β	per capita successful transmission rate	variable
γ	recovery rate	$1/18$ (day^{-1})
α	rate of infection (from exposure)	$1/12$ (day^{-1})
ρ	return rate	$1/5$ (day^{-1})

A.1 Spread

Smallpox usually spreads slowly;[4] however, brief exposures have occasionally led to infection. One index case led to ≥ 12 people being infected (outbreak in Yugoslavia, 1972). Nomad communities in the Horn of Africa had spontaneous termination of the disease in $> 98\%$ of recorded outbreaks. However, if transmission was maintained, it could continue for many weeks in these small populations. Airborne infection over short distances did sometimes occur. Several hospital outbreaks in the 1960s and 1970s were thought to have had an airborne source and transmission pathway. Several circumstances favoured the airborne transmission of variola virus: (*i*) the index patient had a densely confluent rash and a severe cough, and (*ii*) the relative humidity in the hospital was low, a situation that promotes the survival of the

virus.[5] The clothing and bed-linen of smallpox patients are heavily contaminated with the virus. These objects sometimes served as a secondary source of cases. There are several documented instances of laundry personnel working near a hospital or hotel where a smallpox victim had been nursed, who became infected.[6]

A.1.1 Spatial Distribution

Pre-colonial Australia had a complex network of economic and cultural connections. These connections have been categorised into two main types of population groupings: the local group or 'band' population, and the culture-area population based on drainage divisions.[7] The boundaries of these groups are not set and subject to fluctuations and long-term change. We attempted to capture some of this structure in our movement model by grouping populations in cultural or language groups and within the 17 larger groupings of patches corresponding to the geographical (drainage) divisions. This is of course subject to debate and different authors have attempted to group cultural areas into several proposed classifications.[8–12]

Historic accounts of smallpox Smallpox was endemic to Europe and Asia for centuries before the age of exploration spread it to new populations. It is therefore difficult to gauge the disease's impact on previously unexposed populations, even though several historical examples provide clues. A smallpox epidemic spread through Japan in 737, perhaps the first to appear in the country. Detailed records of rice loans suggest that from 735 to 737, some areas experienced mortality $> 60\%$ or even $> 70\%$.[13] Subsequent smallpox epidemics were brought to Japan over the next several hundred years until the disease became endemic at the end of the Second Millennium. In North and South America, it is suspected that as much as 95% of the Indigenous population died due to the introduction of Old World diseases — smallpox is thought to be the main culprit.[14,15] In 1720, smallpox first appeared among the 200 inhabitants of Foula Island, north of Scotland, killing 90% of the island's population.[16] Historians have pieced together various accounts from colonists, explorers, and military personnel to construct an approximate timeline of possible events. Much relies on a few first-hand accounts by colonists who witnessed pockmarked faces, and oral histories describing the local impact of sickness. A series of later, more restrained outbreaks in the late 19$^{\text{th}}$ Century predominantly affected isolated port cities and were mostly confined to European colonists.[17]

A.2 Basic reproduction number

The basic reproduction number $\mathcal{R}_0$ is the average number of secondary cases from a single source. Thus, $\mathcal{R}_0$ affects the speed of transmission and thus the spatial distribution of epidemics. Because it is difficult to estimate the true transmission rate, we used a general estimate of the basic reproduction number from historical outbreaks of smallpox. Based on several historical outbreaks, Gani and Leach[18] estimated that the average value for smallpox was between 3.5 and 6, although it was occasionally much higher in densely populated European cities such as London or Paris. Estimates based on well-documented epidemics where smallpox was endemic for centuries do not necessarily capture the dynamics in the epidemiologically naïve Indigenous population of Australia. It is therefore possible that the $\mathcal{R}_0$ we used are downwardly biased. For completeness, we explored a range of $\mathcal{R}_0$ to account for uncertainty in the predicted rate of geographic spread and total mortality.

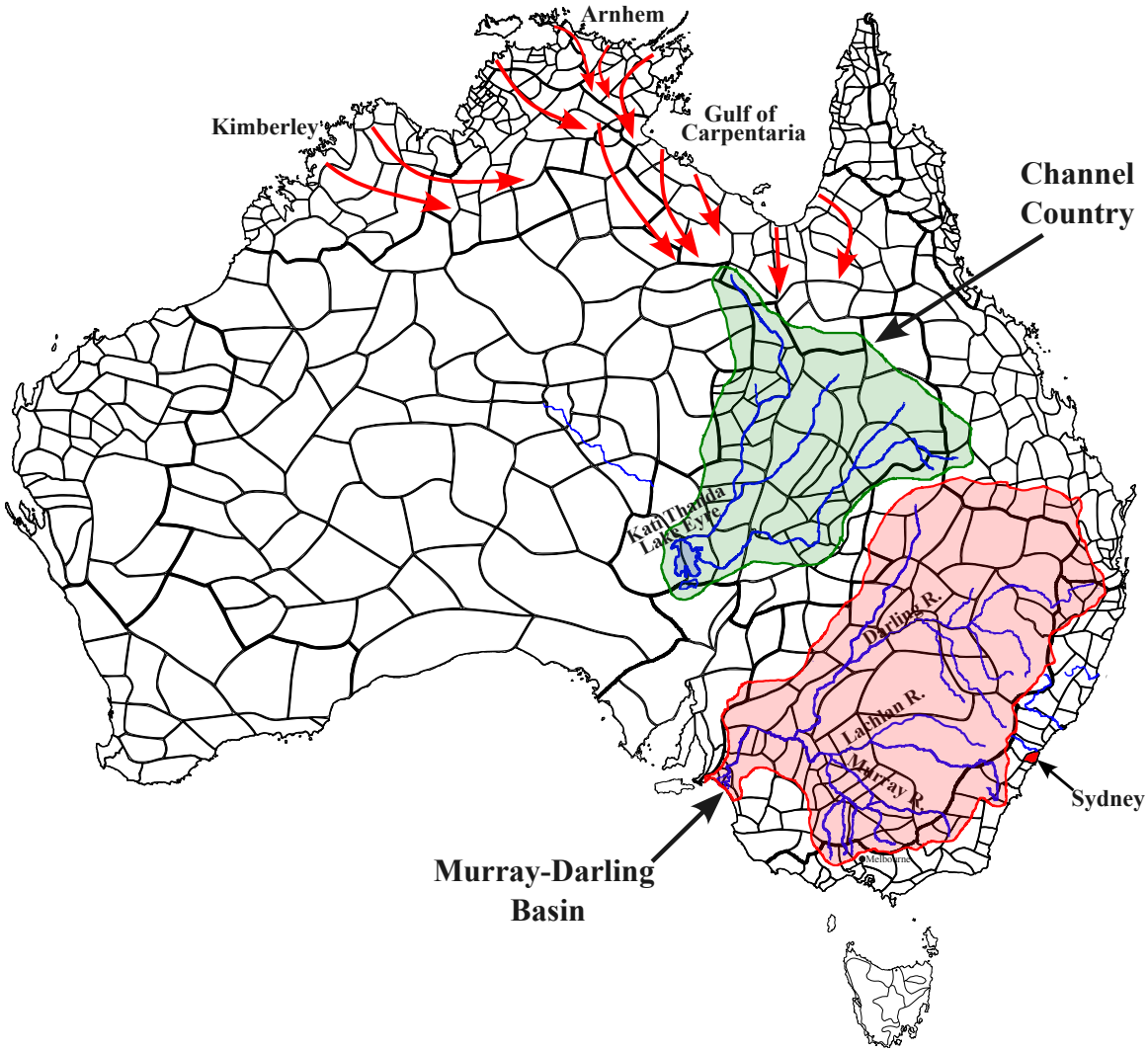
We applied a constant $\mathcal{R}_0$ to all patches to estimate the average number of successful transmissions between individuals. In the absence of a natural birth or death rate, for a multi-patch system, $\mathcal{R}_0 = \beta/\gamma$.[19] Because we can be confident of the mean duration of the infectious period, this produces an estimate for the mean successful contact rate at $\beta \approx \mathcal{R}_0\gamma$.

A.3 Makassan introduction hypothesis

The hypothesis that smallpox first arrived in Australia from Indonesia, probably with Makassan traders and fishers, was originally proposed starting in the early 20th Century.[20–23] This hypothesis proposed that smallpox was spread to the northern coast of Australia through contact between Makassan visitors and local Aboriginal people and once introduced, smallpox spread southward along Aboriginal trade routes and eventually appeared near Sydney when members of the First Fleet identified its first victims.

Proponents of this hypothesis point to several accounts of smallpox (or an unnamed disease) that seemed to point chronologically to a spread south. There are many artifacts linking Makassan fishers to Australia where they routinely formed large camps along the coast to process trepang (sea cucumbers). This exposure to the coasts of Australia led to substantial contacts between these fishers and local Aboriginal people, with some Australians travelling back to Indonesia with their contacts.[24] Cumpston[22] was less convinced of the Makassan hypothesis for the first outbreak of 1789, but argued that it was the best explanation for the outbreaks of the 1860s. Butlin[25] was one of the first modern scholars to reject the Makassan hypothesis, arguing that smallpox must have been released in 1789 near Sydney. He noted that the time required to travel from Sulawesi of more than seven to eight weeks would have made the chance of transmission limited (low contact between Aboriginal people and the Makassan fishers, lack of clothing as a carrier, and the virus is destroyed in salt water). In his view, the exposed Makassans were either dead or fully recovered by the time they reached the Gulf of Carpentaria.

Figure S1: Makassan introduction hypothesis. Campbell[26] and others[20,21] proposed that Makassan contact in Arnhem and the Gulf of Carpentaria brought smallpox to Australia in the late 1780s. They contend that smallpox then spread through the continent via "chains of connections", such as major gatherings, down-the-line exchanges, and continent-scale trade networks. Proponents argue that southward movement through Channel Country and into the Murray-Darling river system could have provided transmission across much of the eastern half of the continent.



B Mathematical details

B.0.1 Mobility model

Residence patches represent each of the distinct language and cultural groups on the continent.[27] To model the links between each of these language groups, we constructed a network of M nodes corresponding to the approximate location of each distinct group. The edges between nodes represent the nearest-neighbour connections that we assigned a distance of 1.

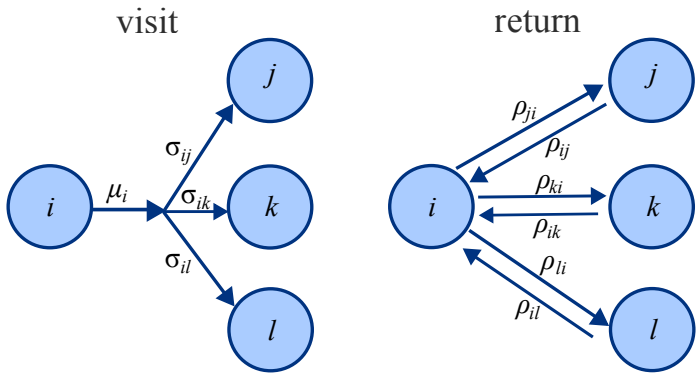
The traditional mathematical approach to compartmental models of epidemics is to assume that patches are cities with equal area (and often equal population as well). This will lead to different dynamics with a mostly rural, hunter-gatherer population who travel widely within their residency patch, and therefore might interact infrequently with other members of the same residency patch.

We therefore assumed a patchy environment in which individuals are in discrete patches of varying area and initial population size. Patches are arranged on a grid with fixed positions,

with edges that represent a path between them. We focused on a collection of patch configurations that are defined by the edges that link the various patches (see Appendix D). Each configuration represents a possible geographic situation where links between regions might be restricted by natural barriers such as rivers and mountains, and others that are more likely due to the proximity of resources and ease of travel.

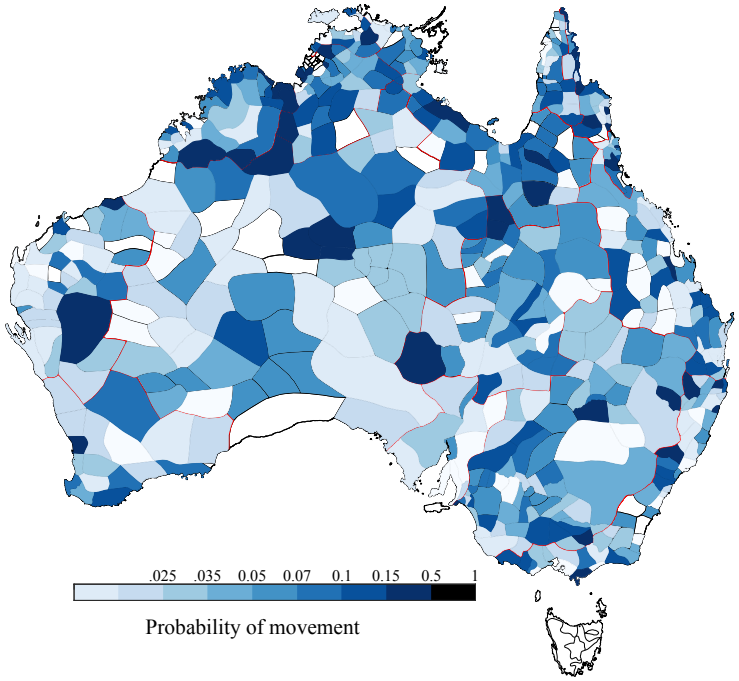
We focused on short-term travel and not on permanent migration because of the rapid transmission behaviour of smallpox. Thus, each individual is assigned a residency patch and returns to this patch at the end of each visit before leaving home again for an additional visit. Both the rate individuals visit neighbouring patches and the rate at which they return home are fixed.

Figure S2: Movement model. At each unit of time, residents of patch i move out of the patch to neighbouring patches with probability μ_i. The probability of visiting a neighbouring patch is given by a distribution over all neighbouring patches, so the probability of a sojourn to patch j from patch i is σ_{ij} . In the same time step, each visitor returns to the home patch with probability ρ. Thus, residents of patch j return to patch i with probability ρ_{ij} and similarly residents of patch i in patch j return at rate ρ_{ji}.



The probability of movement is defined as the product $\mu\sigma$, meaning that each individual has a probability of moving to a particular patch (Fig. S2). For simplicity, we assume that individuals can only move to neighbouring patches because of the short duration of smallpox outbreaks, and do so with a probability determined from previous work.[28,29] This comes from the likelihood of moving along 'superhighways'[29] of pre-agropastoralist travel, compiled from the relative distance to water and other resources. An image of the probability assigned to each patch is shown in Figure S3, with darker patches more likely destinations for visitors than lighter-coloured patches.

Figure S3: Probability of geographic movement. Each patch in our model is assigned a probability that corresponds to the chance that a neighbouring visitor travels there. Darker patches are sinks and lighter patches are sources. We obtained probabilities from previous work.[28,29]



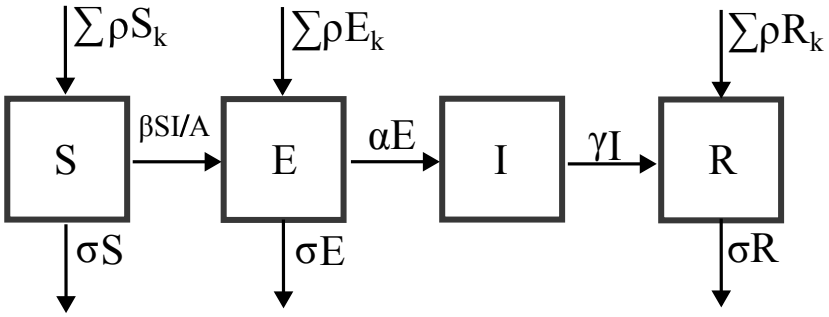
B.1 Epidemiological model

In each of the M patches, we formulate an S-E-I-R model with one type of mobility as described in the previous section. Let S_{ik} and I_{ik} denote the number of susceptible and infectious individuals resident in location i who are present in location k at time t. We modelled disease transmission using a density-dependent incidence, which captures more of the spatial heterogeneity in 1788 Australia. Thus, the infection rate of individuals from patch i currently in patch k is given by

$$\sum_{j=1}^{M} \frac{S_{ik} I_{jk}}{A_j}, \tag{1}$$

where β is the fraction of successful transmissions of smallpox. In our model, we assumed β is constant. Figure S4 illustrates the within-patch compartmental model.

Figure S4: Within-patch model. All individuals in each patch are in one of the four disease compartments: susceptible, exposed, infectious, or removed (recovered), also typically denoted 'S-E-I-R'. Smallpox victims became immune to the disease if they survived. Healthy individuals moved from patch to patch, thus all S, E, and R include both visitors and residents of the patch. Infectious individuals are assumed too ill to move and so once infectious, stay in their home patch. Other parameters: ρ = return rate for visitors out of a patch, σ = rate at which individuals leave the patch for a visit elsewhere, β = fraction of successful transmissions of smallpox, α = rate at which exposed individuals become infectious, and γ = rate infectious individuals are removed from the infectious system (recovered or dead).



B.2 Markov chain model

We opted for a stochastic model modified from the above system of ordinary differential equations considering the size of the region we were modelling. For a few infectious individuals, the predictions of ordinary differential equations can be misleading. In particular, if $\mathcal{R}_0 > 1$, there is a possibility that in a Markov chain model, infectious individuals die or recover before an outbreak occurs. Thus, a Markov chain model is more realistic. To formulate the Markov chain, let the state of system be specified by the vector:

$$X(t) = \Big(S_1(t), E_1(t), I_1(t), R_1(t), ..., S_M(t), E_M(t), I_M(t), R_M(t)\Big) \tag{2}$$

that denotes a discrete-valued random vector with elements $S_{ij}(t)$, $E_{ij}(t)$, $I_{ij}(t)$, $R_{ij}(t)$ for $i = 1, \ldots M$. Thus, the disease compartments are formulated as MM matrices, with the main diagonal representing residents and off-diagonal elements the visitors.

Within each patch, we simulate a standard S-E-I-R compartmental model of disease spread with no age or sex structure; therefore, all individuals are in one of the four states of disease. For simplicity, we consider all those who have transitioned out of the I compartment as 'removed' (R) because they are no longer able to spread the disease and so do not affect the disease dynamics. To estimate mortality, we computed a percentage of the population of R. Figure S5 shows an illustration of the within-patch dynamics. Each individual is categorised in one of the disease compartments with new visitors arriving and mixing with the occupants of the patch while others return home. We assumed each patch is well-mixed and each of the susceptible individuals in a patch are equally likely to make a successful contact with an infectious person.

C Additional supporting results

In this section, we provide some additional, more detailed results to support our conclusions. We show the mean final size of epidemics starting from Sydney and various points along the northern coast. We also provide some S-E-I-R curves for some simulated epidemics and show how the epidemic propagates throughout surrounding patches. Finally, we include detailed tables with average results from each individual patch considered.

C.1 Final size of epidemics

Figure S5 shows the median total daily infections for various values of $\mathcal{R}_0$ in each of the northern coastal regions. The average duration for an epidemic starting in the north is between approximately 2 and 8 years, although there are individual realisations when the disease lingers for much longer.

Figure S5: Average total daily infections for epidemics starting in each of the four northern coastal regions considered in this study. Arnhem: 19 patches, Gulf: 14 patches, Kimberley: 14 patches, North: 10 patches.

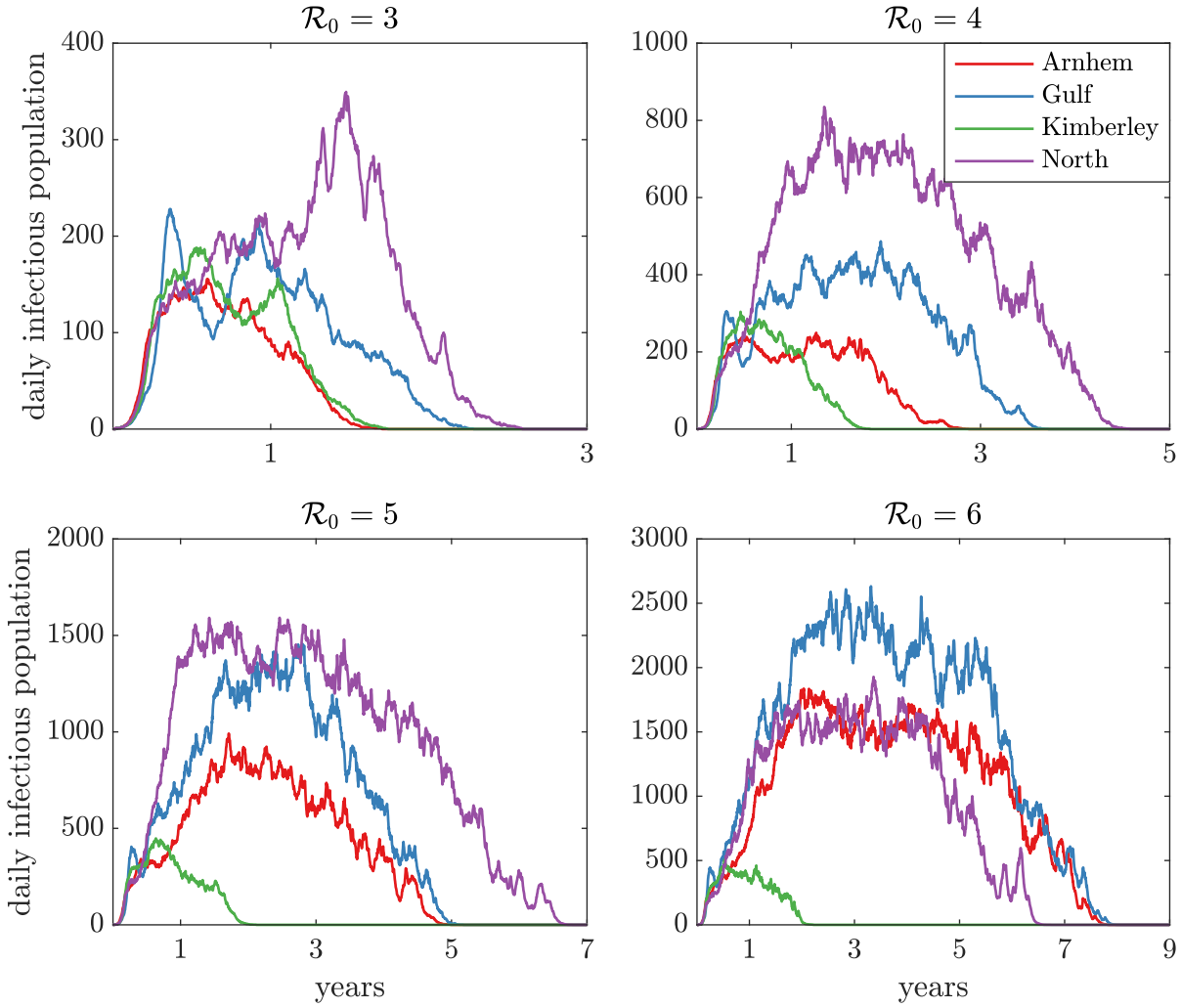
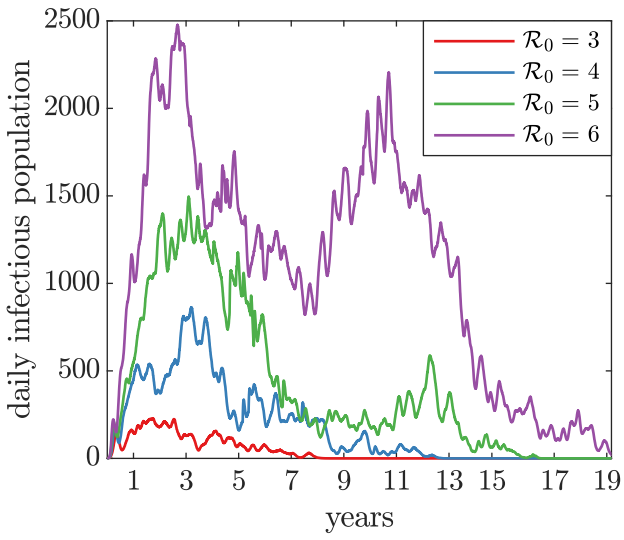


Figure S6 shows the mean number of daily infections for various values of $\mathcal{R}_0$ corresponding to an epidemic originating in Sydney.

Figure S6: Average total daily infections for simulations starting in Sydney with a single exposure.



C.2 Example simulated epidemic

A single epidemic is simulated starting with a single exposure in patch 31, the sight of the Wurrwurrwuy stone arrangements mentioned in the main text. This is only one likely entry point considered in the Makassan hypothesis. Figure S7 shows the S-E-I-R curves within the first few patches of the epidemic as it spreads outward from its initial starting point.

Figure S7: S-E-I-R curves for an epidemic starting in patch 31 (site of Wurrwurrwuy stone arrangements). (a) Curves for the entry patch, patch 31. (b) Curves showing the progression of the epidemic throughout two other neighbouring patches, patch 32 and patch 46. Each simulation is stopped when the saturation point is achieved. Deaths are not included and are thus a portion of the R (removed/recovered) population.

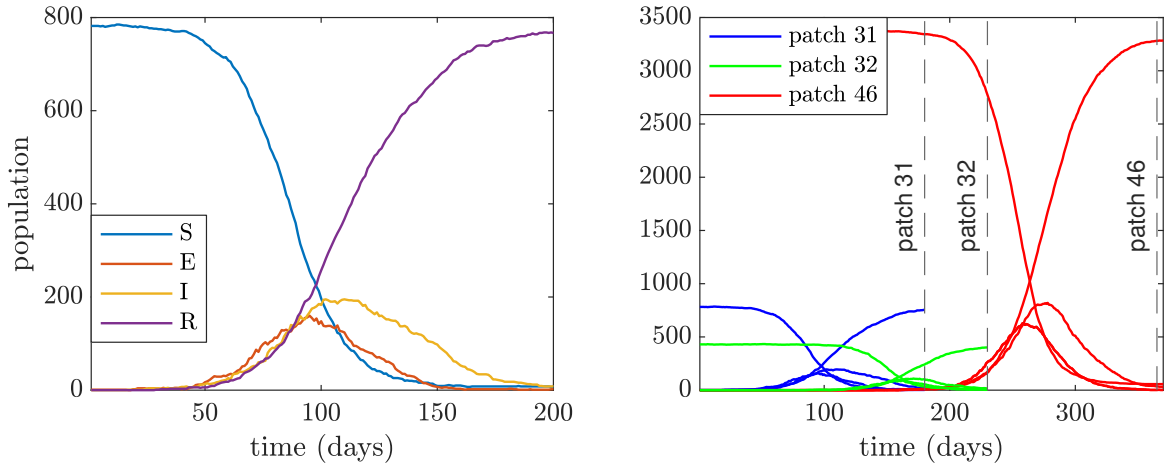


Figure S8 shows the progression of a simulated epidemic as it moves through time and space. This is an example in which a single exposed individual started in patch 31 over the course of the first 1000 days of the epidemic. The movement between nearby patches causes the disease to spread and a wave of peaks cascades from the initial exposed patch.

Figure S8: Simulated epidemic starting from an exposure in patch 31 for the first 1000 days. Curves represent the total daily infections per day within individual patches and vertical dashed lines show the time when each patch achieved the maximum daily number of infections (numbers indicate the patch).

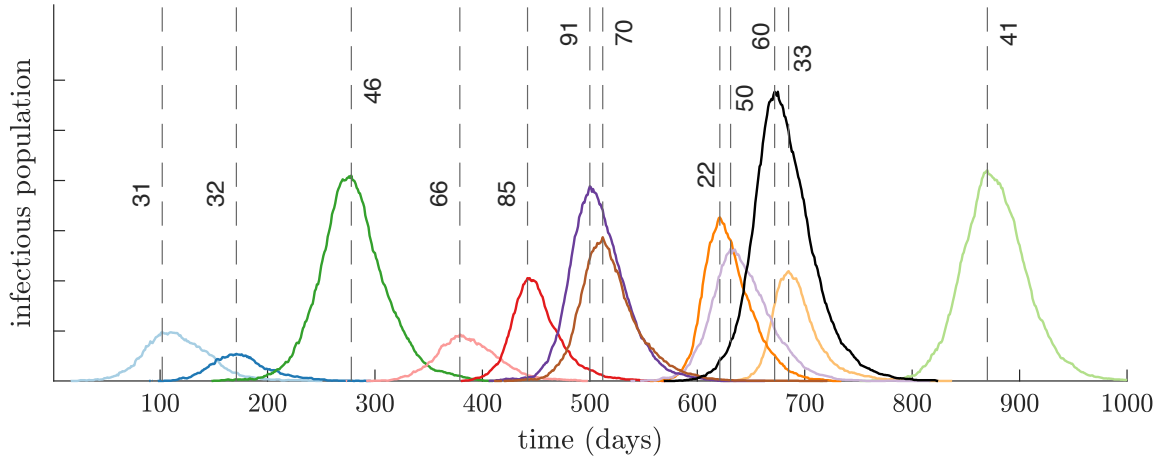
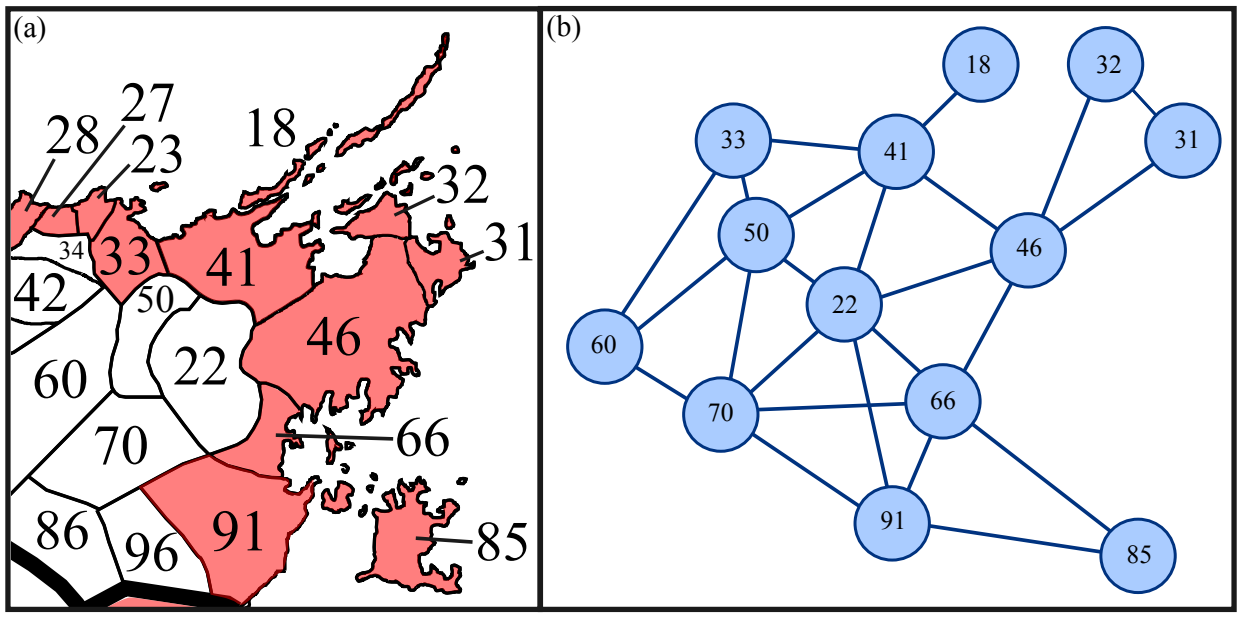


Figure S9: Closeup view of the epidemic region. (a) The geographic map of the affected region of Arnhem Land, including all of the patches where smallpox was present. (b) A network representation of connected patches that we used to determine the progression of the epidemic originating with a single exposure in patch 31. Edges represent possible two-way movement between patches which are represented as nodes.



The simulated epidemic lasts < 4 years; however, there are some simulations showing an epidemic lingering for as much as 10 years, generated by movement between patches that provides the virus with a new population of susceptible individuals (Fig. S9).

C.3 Individual Patch Results

Table S4 lists the average results for epidemics originating in single patches within each northern region explored in the main text. Full results are shown for several fixed values of $\mathcal{R}_0$.

Table S4: Average results for epidemics with various fixed $\mathcal{R}_0$ grouped into the four regions considered in this study. In each patch, we simulated multiple realisations and combined them to determine the expected behaviour of an epidemic of smallpox. We assumed that mortality is based on a death rate of 60% for those infected with the disease. The most likely spatial extent of an epidemic is tabulated in the final column of the table (number of total patches affected). To obtain these numbers, we computed the maximum probability that at least one infectious individual entered a given patch.

Origin coastal region	I_M	I_{Mdn}	SE_I	$dths_{Mdn}$ (@ 60%)	SE_{dths}	dur_M (days)	dur_{Mdn} (days)	SE_{dur}	Greatest patch spread ($p \geq 0.5$)
Arnhem	**19 coastal patches**								
$\mathcal{R}_0 = 3$	27,058	5,336	1,833	3,202	1,100	1,096	622	47.1	24
$\mathcal{R}_0 = 4$	93,360	22,574	4,911	13,544	2,946	1,610	1,043	57.4	41
$\mathcal{R}_0 = 5$	261,140	124,637	15,566	74,782	9,940	2,277	1,820	95.8	73
$\mathcal{R}_0 = 6$	608,540	381,219	28,666	228,731	17,200	2,874	2,861	105.6	146
Gulf	**14 coastal patches**								
$\mathcal{R}_0 = 3$	27,0838	9,753	1,968	5,852	1,181	1,161	827.5	55	16
$\mathcal{R}_0 = 4$	107,570	46,196	6,922	27,718	4,153	1,609	1,350	68.6	31
$\mathcal{R}_0 = 5$	277,160	144,202	18,040	86,521	10,824	2,149	1,888	100.7	131
$\mathcal{R}_0 = 6$	616,490	539,632	32,480	323,780	19,488	2,819	2,917	119	229
North	**10 coastal patches**								
$\mathcal{R}_0 = 3$	34,484	21,797	2,565	13,078	1,539	1,171	945	60.7	14
$\mathcal{R}_0 = 4$	125,880	84,232	8,791	50,539	5,275	1,902	1,671	93.76	49
$\mathcal{R}_0 = 5$	353,980	239,995	22,929	143,997	13,757	2,683	2,440	130.5	98
$\mathcal{R}_0 = 6$	651,440	531,820	36,311	319,092	21,787	3,076	3,464	137.6	253
Kimberley	**14 coastal patches**								
$\mathcal{R}_0 = 3$	13,588	6,243	1,328	3,746	797	864	633	42.2	16
$\mathcal{R}_0 = 4$	61,495	9,426	5,541	5,655	3,324	1,238	668	64.1	17
$\mathcal{R}_0 = 5$	204,930	20,102	18,883	12,061	11,330	1,804	774	109.5	21
$\mathcal{R}_0 = 6$	507,110	104,768	33,315	62,861	19,989	2,493	1,375	126.1	92

The full results for individual patches are displayed in Tables S5,S6,S7 and S8. The corresponding number of total daily infections are shown for each realisation in Figures S10, S11, S12 and S13. Each figure is scaled the same to compare the effect of increasing $\mathcal{R}_0$, including the wide variance. Mean and median curves are included for comparison.

Figure S10: Total daily infections for entry points along the coast of Arnhem Land with the four $\mathcal{R}_0$ values considered in this study. All simulations are shown in grey along with the mean and median curves in blue and red. In each patch, we simulated multiple realisations (> 25). We use the same scale for each panel to illustrate the relative size (number of infections) of each simulated epidemic.

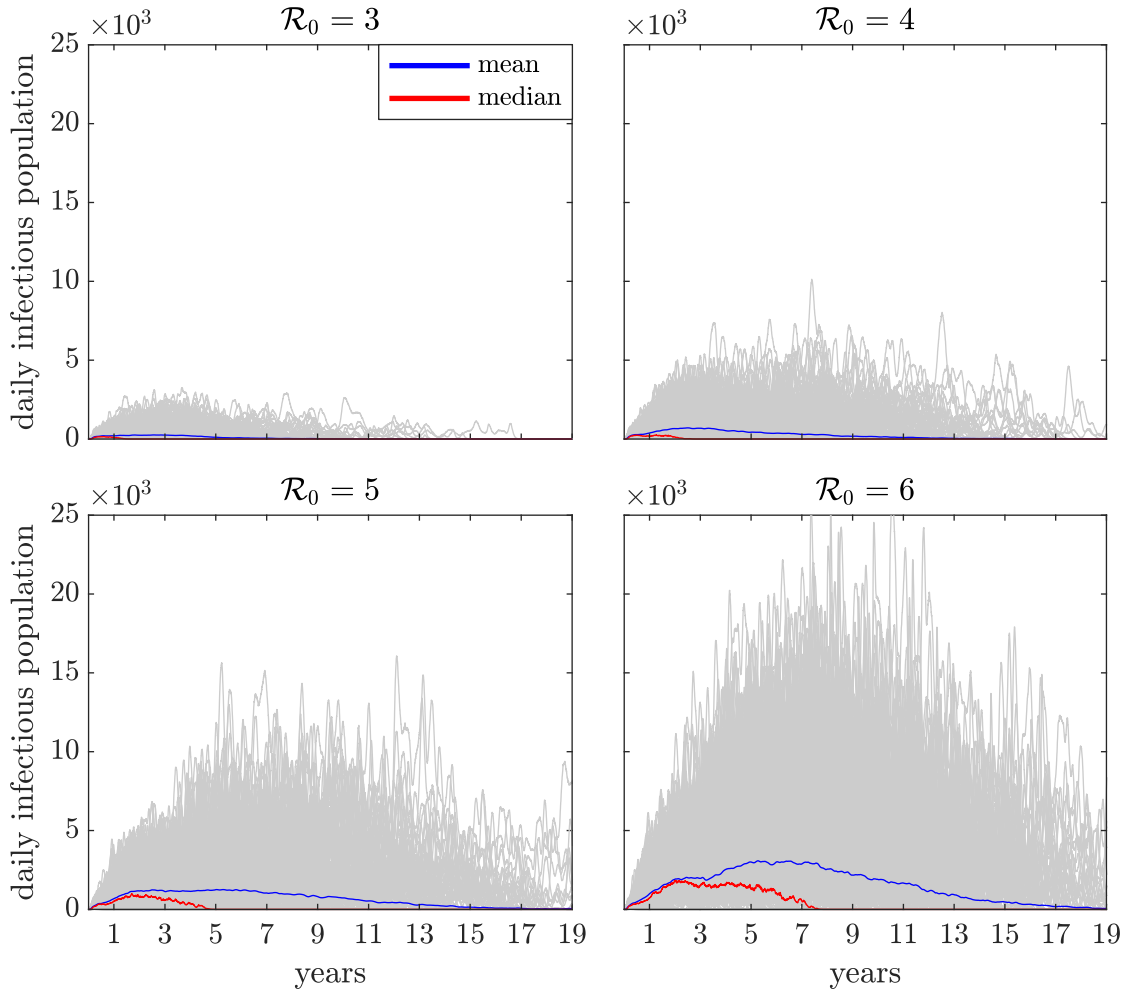


Table S5: Arnhem patches for a fixed $\mathcal{R}_0 = 6$. Details of simulations in each individual patch. The most likely spatial extent of an epidemic is tabulated in the final column of the table (number of total patches affected). To obtain these numbers, we computed the maximum probability that at least one infectious individual entered a given patch.

Origin patch	I_M	I_{Mdn}	SE_I	$dths_{Mdn}$ (@ 60%)	SE_{dths}	dur_M (days)	dur_{Mdn} (days)	SE_{dur}	Reach Sydney	≥ 50 I	Most likely patch spread ($p \geq 0.5$)	Total runs
5	513,970	14,462	102,270	8,677	61,362	2,395	651	402.86	6	35	18	38
6	525,800	211,320	113,250	126,792	67,950	2,567	2,026	426	5	30	79	36
7	595,630	538,630	106,840	323,178	64,104	3,000	3,569	431.9	3	28	136	38
8	575,710	247,962	140,610	148,777	84,366	2,781	1,896	560.5	8	22	109	28
9	724,340	544,562	130,006	326,737	78,003	3,127	2,980	452.2	8	27	157	30
15	275,690	1,869	96,634	1,121	57,980	1,516	298	399	1	21	2	38
18	444,570	8,218	118,260	4,931	70,956	2,179	568.5	460.8	3	28	6	36
23	644,180	431,692	117,100	259,015	70,260	3,020	3,216	394	6	32	139	36
24	492,640	198,746	139,310	119,248	83,586	2,596	1,743	576.4	3	19	84	37
27	725,560	839,970	119,020	503,982	71,412	3,573	3,903	449	6	24	183	37
28	859,710	953,293	108,790	571,976	65,274	4,145	4,697		7	30	262	39
31	253,840	1,202	104,620	721	62,772	1,468	312	425.2	1	19	2	38
32	173,230	1,198	105,820	718	63,492	929.7	308	363.1	2	23	2	36
33	1,003,100	1,144,892	126,500	686,935	75,900	4,156	4,857	406.2	11	25	354	39
41	648,030	495,823	122,040	297,494	73,224	3,074	3,394	400.7	6	27	120	36
46	507,890	252,310	142,720	151,386	85,632	2,476	2,475	415.8	1	18	123	36
66	857,210	933,034	121,430	559,820	72,858	3,978	4,009	388.5	6	23	256	39
85	594,250	215,662	131,780	129,397	79,068	2,743	2,103	478.5	7	25	93	36
91	928,370	979,323	120,160	587,594	72,096	4,040	4,502	386	8	26	317	38

Figure S11: Total daily infections for entry points along the coast of the Gulf of Carpentaria with the four $\mathcal{R}_0$ values considered in this study. Results of all simulations are shown in grey. The mean and median are shown in blue and red. In each patch, we simulated multiple realisations (> 25). We use the same scale for each panel to illustrate the relative size of each simulated epidemic.

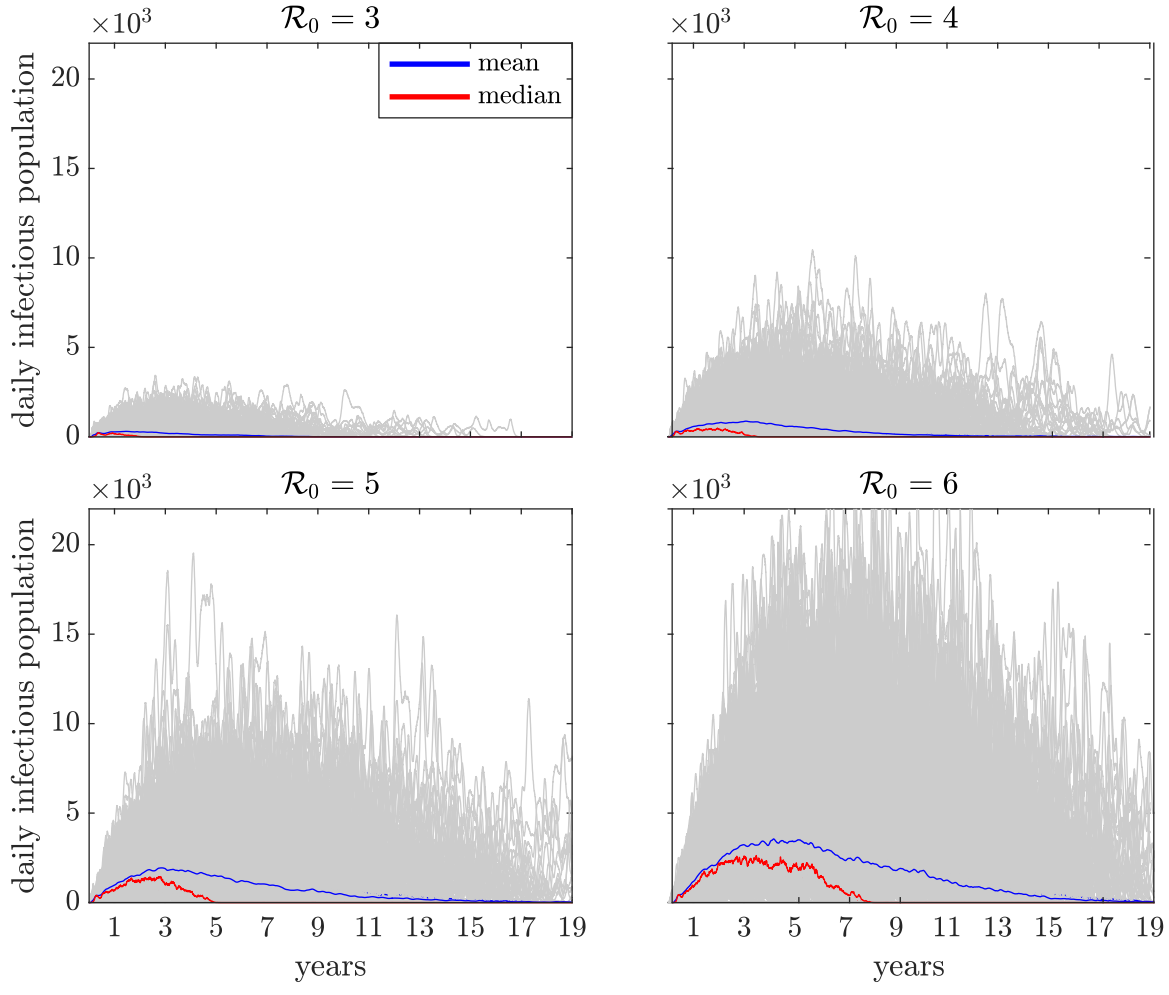


Table S6: Gulf patches for a fixed $\mathcal{R}_0 = 6$. Details of simulations in each individual patch. The most likely spatial extent of an epidemic is tabulated in the final column of the table (number of total patches affected). To obtain these numbers, we computed the maximum probability that at least one infectious individual entered a given patch.

Origin patch	I_M	I_{Mdn}	SE_I	$dths_{Mdn}$ (@ 60%)	SE_{dths}	dur_M (days)	dur_{Mdn} (days)	SE_{dur}	Reach Sydney	≥ 50 I	Most likely patch spread ($p \geq 0.5$)	Total runs
124	876,290	1,025,404	121,960	615,242	73,176	3,708	3,828	410	4	24	298	38
146	1,096,132	1,224,700	122,370	734,820	73,422	4,426	4,818	435.2	7	18	467	37
152	336,000	150	104,880	90	62,928	1,666	174.5	438.5	4	30	1	38
163	840,610	982,007	123,090	589,204	73,854	3,948	4,689	415.6	9	24	299	38
167	908,3802	1,073,700	158,500	644,220	95,100	3,873	4,025	540.3	3	12	328	29
174	200,860	2,678	74,491	1,607	44,695	1,000	226.5	247.9	1	28	2	38
180	637,790	686,147	101,210	411,688	60,726	3,497	3,868	393	9	20	209	34
184	304,540	2,678	82,066	1,609	49,240	1,379	244	288.6	1	33	2	37
190	323,010	150	88,309	90	52,985	1,631	206	373.6	4	27	1	36
207	675,060	725,900	92,715	435,540	55,629	3,390	3,810	316.4	7	28	169	38
209	955,510	859,298	123,450	515,579	74,070	4,263	4,726	361.8	10	20	408	27
222	548,250	395,740	112,930	237,444	67,758	2,167	2,785	435.1	4	27	92	36
225	832,910	917,998	109,610	550,799	65,766	3,450	3,958	263.5	7	22	349	39
229	828,500	663,960	185,900	398,376	111,540	3,567	3,126	515.1	3	13	281	37

Figure S12: Total daily infections for entry points along the Kimberley coast with the four $\mathcal{R}_0$ values considered in this study. Results of all simulations are shown in grey. The mean and median are shown in blue and red. In each patch, we simulated multiple realisations (> 25). We used the same scale for each panel to illustrate the relative size of each simulated epidemic, given a fixed $\mathcal{R}_0$.

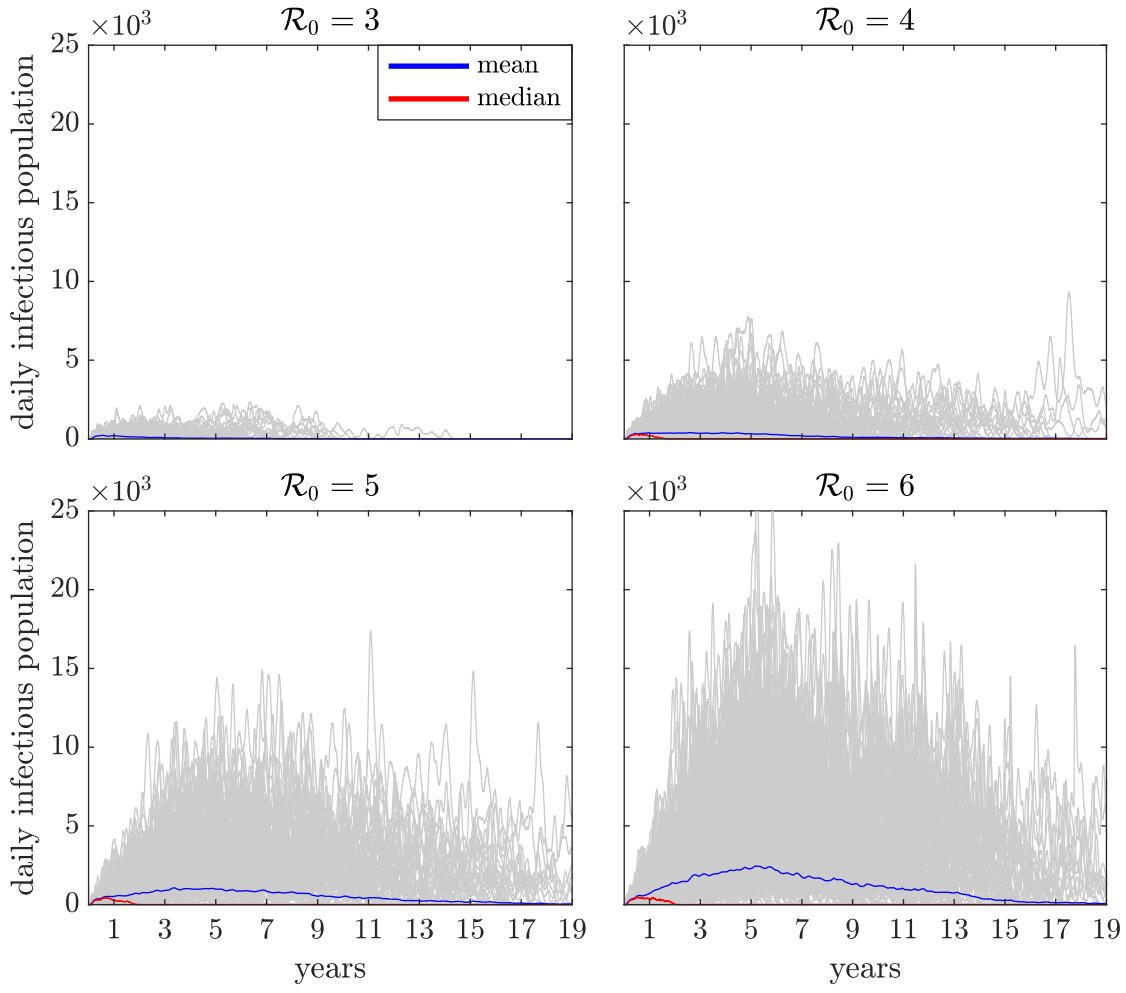


Table S7: Kimberly patches for a fixed $\mathcal{R}_0 = 6$. Details of simulation for each individual patch. The most likely spatial extent of an epidemic is tabulated in the final column of the table (number of total patches affected). To obtain these numbers, we computed the maximum probability that at least one infectious individual entered a given patch.

Origin patch	I_M	I_{Mdn}	SE_I	$dths_{Mdn}$ (@ 60%)	SE_{dths}	dur_M (days)	dur_{Mdn} (days)	SE_{dur}	Reach Sydney	≥ 50 I	Most likely patch spread ($p \geq 0.5$)	Total runs
100	536,030	459,448	118,200	275,669	70,920	2,636	2,005	427.5	2	23	123	39
104	754,020	885,471	144,800	531,283	86,880	3,699	4,057	577.6	7	20	146	37
111	626,980	413,600	163,090	248,160	97,854	2,960	2,672	581.8	3	16	126	35
127	709,460	658,516	126,900	395,110	76,140	3,442	3,828	435.6	4	21	157	36
129	741,850	753,580	125,960	452,148	75,576	3,770	4,573	426.7	7	20	163	36
136	854,710	519,587	142,950	311,752	85,770	3,714	4,053	467.3	10	26	186	36
151	691,550	695,910	129,090	417,546	77,454	3,540	3,686	548.6	2	20	197	39
169	452,660	30,779	105,360	36,468	63,216	2,233	1,366	429.8	1	27	14	36
172	290,400	8,042	87,472	4,825	52,483	1,638	568	330.2	1	31	8	36
176	293,960	13,375	106,580	8,025	63,948	1,518	537	397	3	31	12	37
183	389,780	11,772	113,700	7,063	68,220	1,904	622	382.3	3	25	11	36
197	335,650	8,025	100,040	4,815	60,024	1,838	429	443.5	1	26	7	37
210	373,490	10,840	127,360	6,504	76,416	1,848	474.5	500.9	3	22	11	38
219	330,340	6,786	123,830	4,072	74,298	1,467	415	405.4	2	23	7	37

Figure S13: Total daily infections for entry points along the coast of the North region with the four $\mathcal{R}_0$ values considered in this study. Results of all simulations are shown in grey. The mean and median are shown in blue and red. In each patch, we simulated multiple realisations (> 25). We use the same scale for each panel to illustrate the relative size of each simulated epidemic, given a fixed $\mathcal{R}_0$.

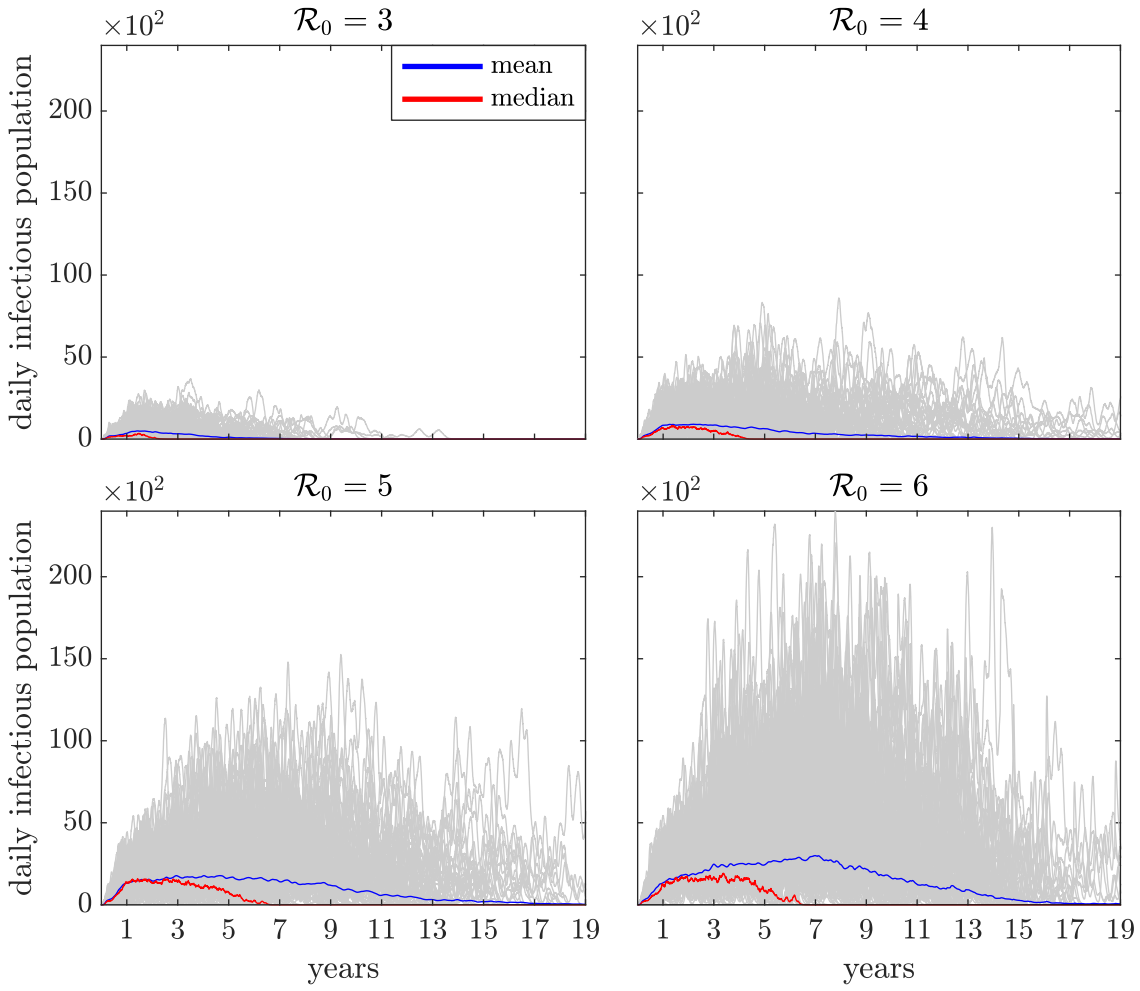


Table S8: North patches for a fixed $\mathcal{R}_0 = 6$. Details of simulation for each individual patch. The most likely spatial extent of an epidemic is tabulated in the final column of the table (number of total patches affected). To obtain these numbers, we computed the maximum probability that at least one infectious individual entered a given patch.

Origin patch	I_M	I_{Mdn}	SE_I	$dths_{Mdn}$ (@ 60%)	SE_{dths}	dur_M (days)	dur_{Mdn} (days)	SE_{dur}	Reach Sydney	≥ 50 I	Most likely patch spread ($p \geq 0.5$)	Total runs
11	529,740	353,565	97,250	212,139	58,350	3,007	3,130	424.6	4	29	113	37
19	1,069,200	1,212,439	88,403	727,463	53,041	4,856	5,017	284.7	7	27	407	39
29	389,640	722	105,400	433	63,240	2,008	194	454.7	3	28	2	35
36	308,120	4,297	108,790	2,578	65,274	1,470	390	399	2	23	4	35
39	599,220	316,860	108,090	190,116	64,854	2,705	2,440	330.7	5	32	130	37
40	818,210	721,170	111,660	432,702	66,996	3,836	3,495	362	6	26	261	38
45	781,270	796,854	109,180	478,112	65,508	3,675	3,750	398.6	8	27	167	38
56	942,130	1,001,600	104,170	600,960	62,502	4,032	4,478	348.6	8	28	359	39
59	775,329	908,770	138,490	545,262	83,094	3,525	4,459	538.5	7	24	218	35
68	301,220	478	85,910	287	51,546	1,623	238.5	370.1	3	28	2	36

Figure S14: Maximum probability of spread. For each patch, the maximum probability of at least one infectious individual present is computed. We defined the greatest likely spread from the indicated entry points as the patches with a probability $p \geq 0.5$ over all simulations. Some isolated darker patches are surrounded by lighter patches because individuals enter a given patch from multiple possible directions. For all simulations, $\mathcal{R}_0 = 6$ and movement parameters are fixed as defined in the main text.

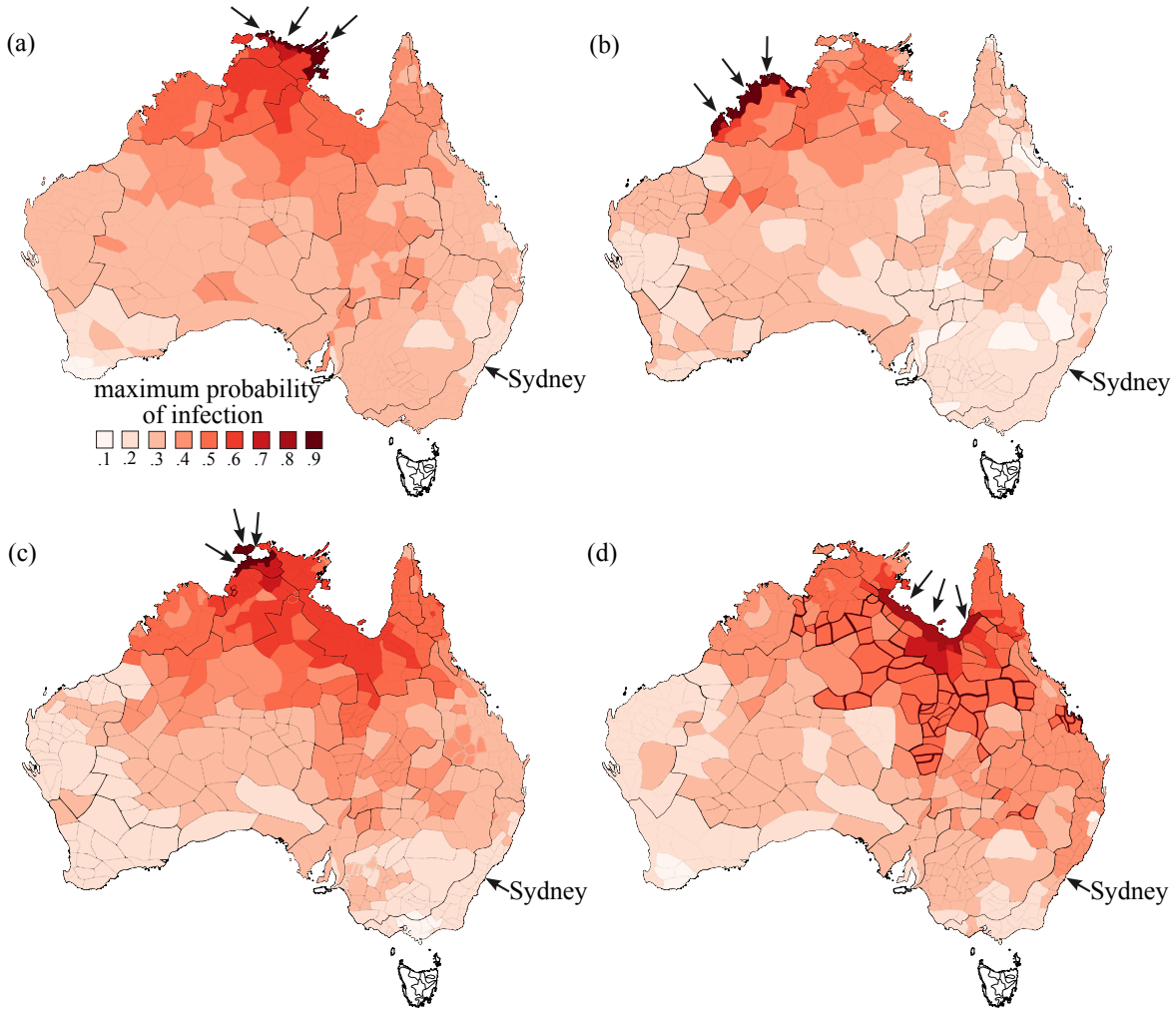
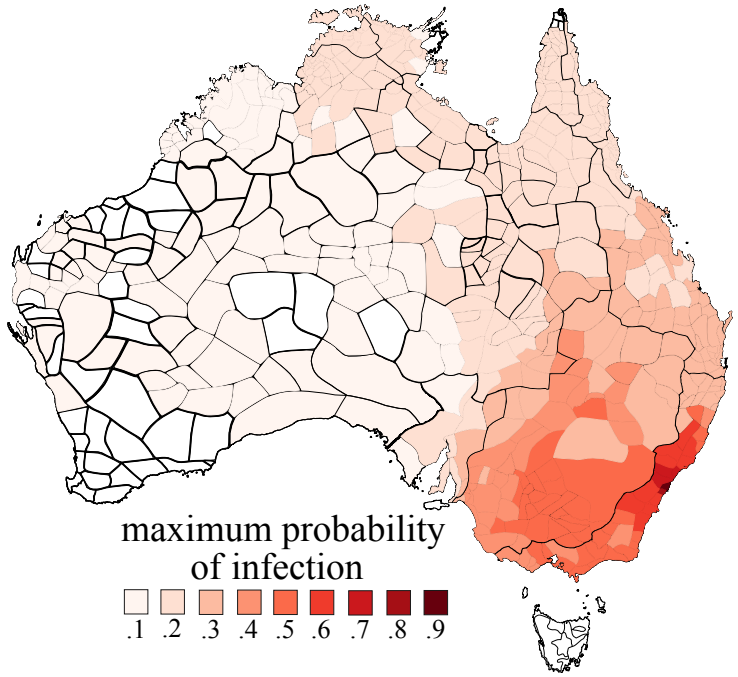


Table S9: Average results for epidemics with various fixed values of $\mathcal{R}_0$ starting with a single exposure in Sydney. In each case, we simulated multiple realisations and combined them to determine the expected behaviour of an epidemic of smallpox. We assumed a mortality of 60% for those infected with the disease. We determined the most likely spatial extent of an epidemic originating in Sydney by considering the probability that at least one infectious individual appeared in a given patch.

$\mathcal{R}_0$	I_M	I_{Mdn}	SE_I	$dths_{Mdn}$ (@ 60%)	SE_{dths}	$dths_{Mdn}$ (yr^{-1})	dur_M (days)	dur_{Mdn} (days)	SE_{dur}	Greatest patch spread ($p \geq 0.5$)
3	9,007	805	2,653	483	1,591.8	690	521.4	255	105.9	3
4	73,077	8,521	20,927	5,113	12,556	3,099	1,351	602	261.3	6
5	168,250	34,064	44,562	20,438	26,737	8,275	1,650	903	288.5	10
6	441,600	223,823	75,704	134,294	45,422	23,981	2,418	2,043	308.1	57

Figure S15: Probability of maximum spread. The probability that a single infectious individual enters a given patch is shown for an epidemic starting in Sydney with a single exposure. $\mathcal{R}_0$ fixed at 6.

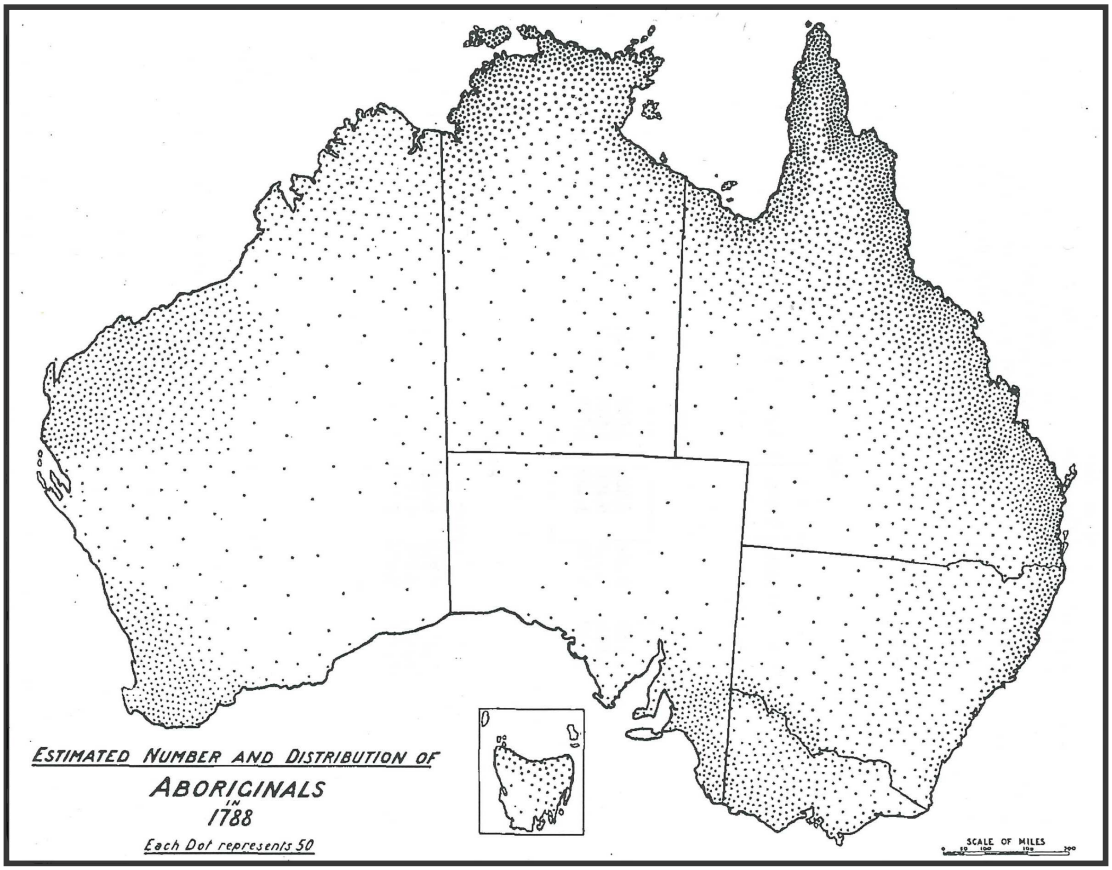


D Population structure

D.1 Population distribution

Little is known of the distribution and size of the human population in pre-1788 Australia. Many estimates have been proposed using different techniques (see[30] for a detailed summary). In 1930, Radcliffe-Brown completed an assessment for the Commonwealth Year Book, which became the standard estimate for much of the 20th Century. That was a systematic study in which several sources for various regions of each state were examined. While he did note that the data were scant and mostly unreliable, he did not attempt to estimate the consequences of disease and concluded the population in 1788 was possibly $>$ 300,000.

Figure S16: Radcliffe-Brown's estimate of the 1788 Australian population. This is based on a total of 300,000 people across the entire continent. Each dot on the map represents 50 individuals. Image reproduced from the Commonwealth Year Book, 1930.



Radcliffe-Brown's[31] survey suggests he considered all his regional population estimates to be minimal, which would allow for a total $> 300{,}000$. Birdsell[32] concluded that there were at least 500 nations, each averaging 500 individuals, although this was a conservative estimate. Tindale[27] later increased the number of groups to 592; however, he reduced the mean patch population to around 450 individuals. More recently, scholars have projected a much larger population. Butlin[25] suggested an aggregate Australian 1788 population of 1.25 million, while Mulvaney and White (1987) suggested a slightly lower population of 750,000. A more detailed assessment of hindcasted carrying capacity, genetic estimates of effective population size, and estimates derived from accumulation rates of archaeological material led Williams et al.[30] to suggest an even higher median population of 2.51 million.

D.2 Language and cultural groups

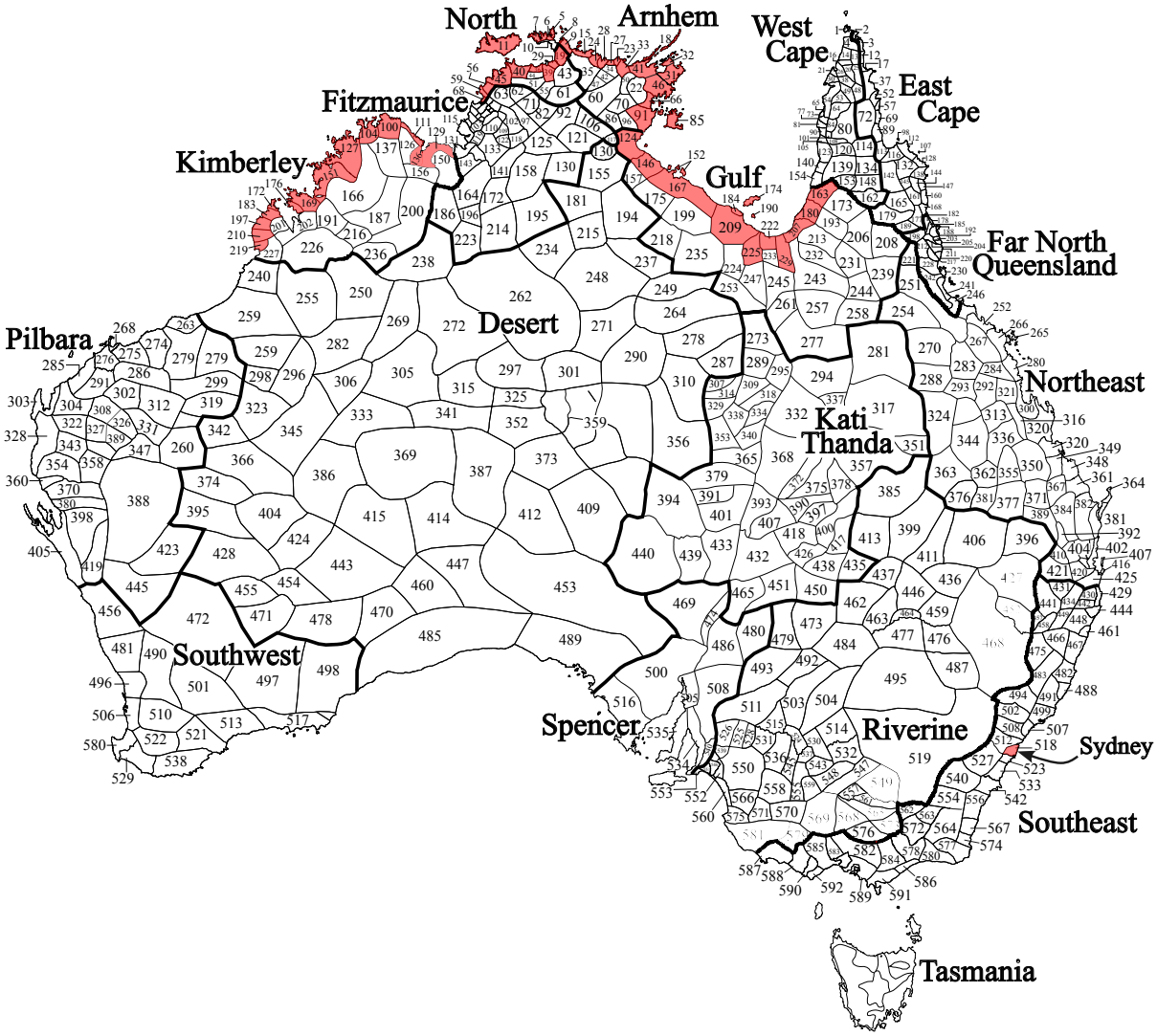
Tindale[33] was the first Australian anthropologist to propose a comprehensive map of the language groups of Aboriginal Australia. Beginning in 1921, he made several ethnographic visits throughout the country, resulting in a map that represented Australia before European invasion. Horton[34] revised the Tindale map to reflect a modern view of Aboriginal Australia in which tribal boundaries were blurred and many of the divisions were combined, which resulted in a map with 390 distinct groups.

Peterson[7] first proposed a division of the continent into 17 distinct cultural areas based on the general knowledge of linguistic and cultural differences. Each region presents a unique challenge for disease to spread. The smaller, more densely packed regions of the north likely supported larger populations with more contact between individuals, where the disease most

likely spread rapidly. Conversely, the large desert region had a much lower population density, and required individuals to travel greater distances, making it more difficult for disease to spread and remain active.

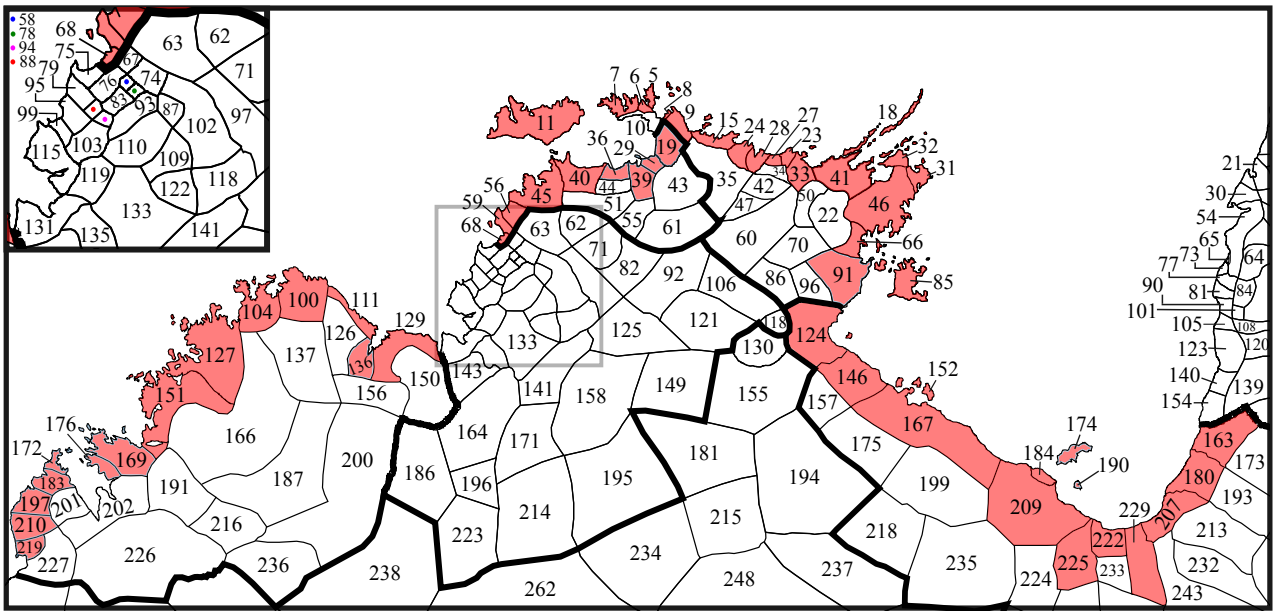
The map we used (Fig. S9) includes all 592 patches and regional divisions. Each patch corresponds to a unique cultural group that we assigned a number and gave an approximate area.[27] Dark lines represent approximate regional divisions and lighter black lines the cultural group. A complete list of all 592 cultural groups and the corresponding patch numbers with estimated populations, approximate country size, and the number of nearest neighbours is included in Table S10. The population sizes are based on an estimate of around 2 million people in 1788.[30]

Figure S17: Tindale Map of Australia. Each of the 592 patches on mainland Australia is labelled with a number. Red patches are all those for which we consider exposed individuals. Patches along the north coast are grouped into four distinct geographical regions: Kimberly, North, Arnhem, and Gulf.



A closeup of the Makassan entry points we considered in red (Fig. S18) span four distinct geographic regions. In total, we simulated epidemics originating from 43 distinct patches.

Figure S18: North coastal region of Australia. All of the 42 separate patches we considered in the simulation are shown in red. Numbers signify individual patches.



D.3 Network Connections

Figures S19, S20, S21, and S22 show a detailed map of each northern coastal region with the corresponding network structure. Nodes represent each patch and lines represent possible connections between patches. For simplicity, we assume that the rate of movement between patches connected by a line segment is the same for each patch.

Figure S19: View of the Arnhem Land coast and the corresponding network structure. (a) A view of the coastal region with patches labelled. (b) A network view of the coast with corresponding patches represented as nodes and lines representing the connections where people can move between patches for travel. All connections are bidirectional and the rate of movement out and return in are the same.

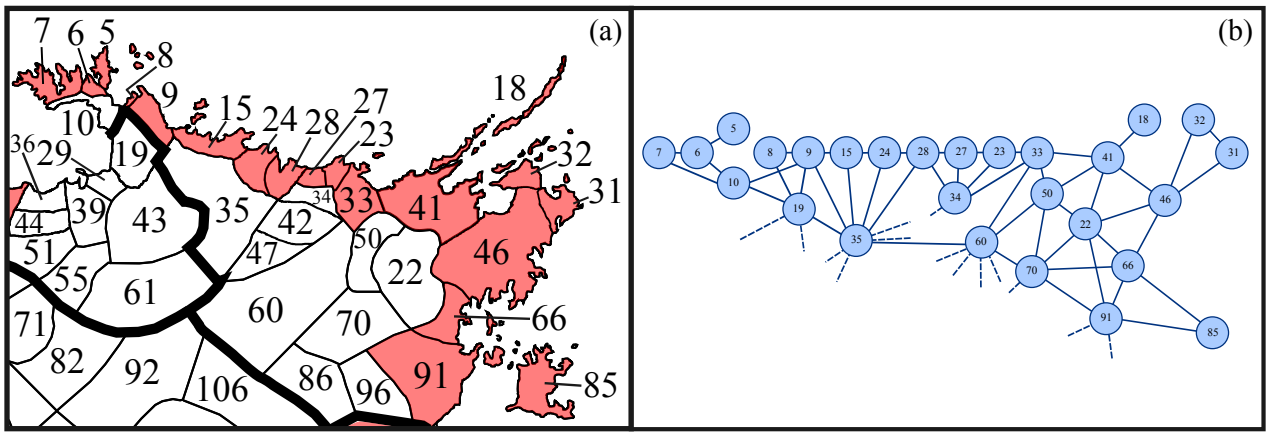


Figure S20: View of the coast of the region called North and the corresponding network structure. (a) A view of the coastal region with patches labelled. (b) A network view of the coast with corresponding patches represented as nodes and lines representing the connections where people can move between patches for travel. All connections are bidirectional and the rate of movement out and return in are the same.

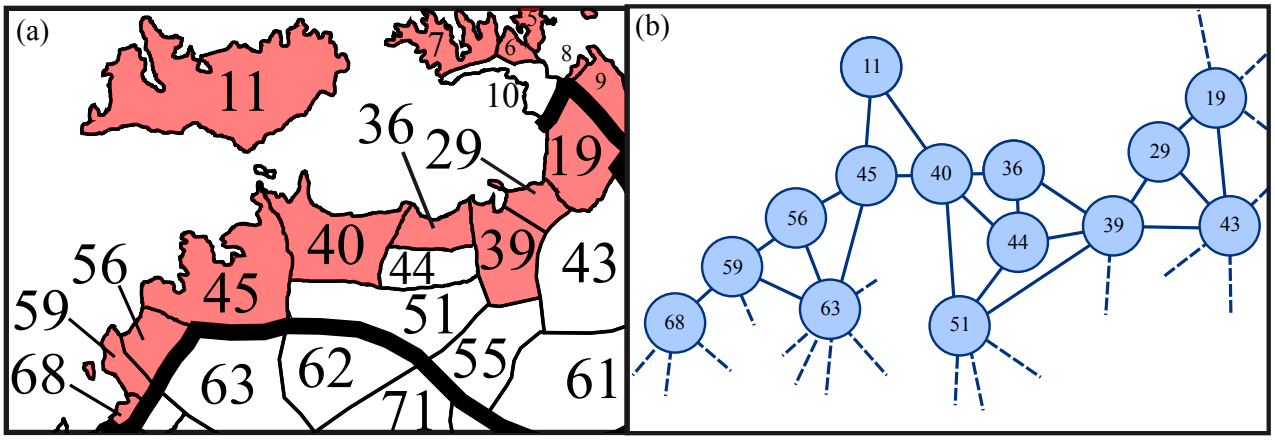


Figure S21: View of the Kimberley coast and the corresponding network structure. (a) A view of the coastal region with patches labelled. (b) A network view of the coast with corresponding patches represented as nodes and lines representing the connections where people can move between patches for travel. All connections are bidirectional and the rate of movement out and return in are the same.

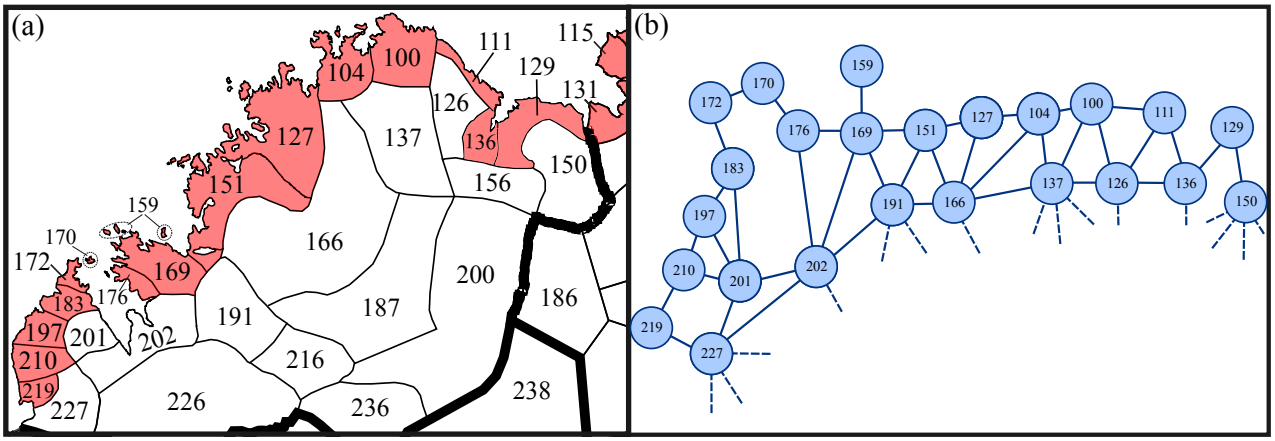
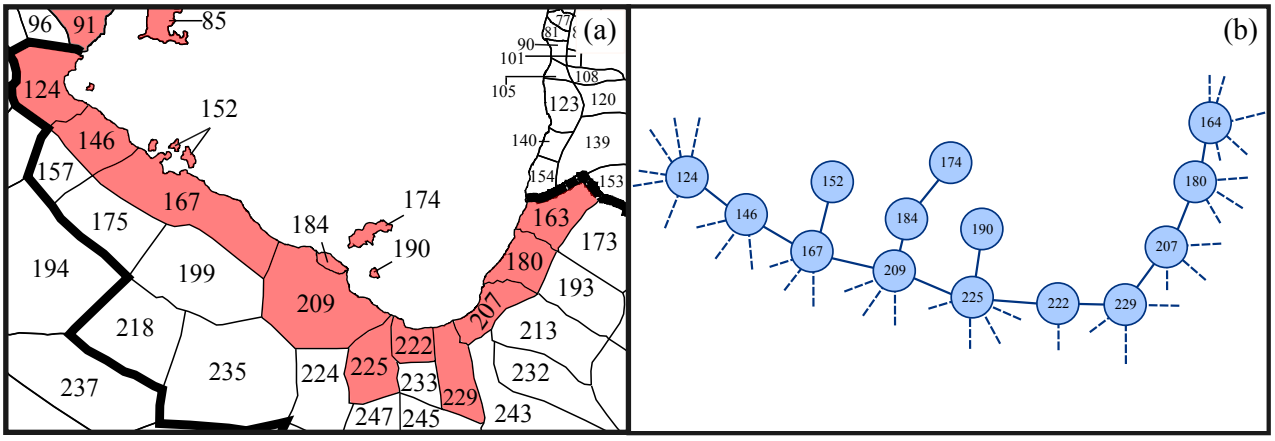


Figure S22: View of the Gulf coast and the corresponding network structure. (a) A view of the coastal region with patches labelled. (b) A network view of the coast with corresponding patches represented as nodes and lines representing the connections where people can move between patches for travel. All connections are bidirectional and the rate of movement out and return in are the same.



D.4 Full list of patches

We provide a complete list of all language and cultural groups that we included in Table S10. We assigned patch numbers based on approximate geographic positions[27] and numbered them from top to bottom and right to left. We based the projected 1788 populations on Williams et al.[30] We used the Williams et al.[30] values along with the original Tindale estimates of the mean total area each group occupied to determine the population densities of each patch. Also included in the table are the number of nearest-neighbouring patches, which for simplicity are the only patches to which visitors travel from their home patch. This gives an indication of how fast an epidemic might have spread with exposed individuals in a given patch. We grouped all patches by the 17 geographic regions (Fig. **??**).

Table S10: Full list of all 592 mainland Australian patches. The projected area and pre-colonial population are included along with the population density and total number of nearest neighbour patches (patches residents can visit).

patch	group	region	modern st/terr.	location (lat/long)	area (km^2)	estimated population	population density	nearest neighbours
1	Kaurareg	West Cape	QLD	(-10.67,142.167)	388	721	1.86	2
2	Djagarago	East Cape	QLD	(-10.83,142.583)	518	721	1.39	3
3	Jathaikana	East Cape	QLD	(-11.167,142.75)	777	721	0.93	3
4	Ankamuti	West Cape	QLD	(-11.167,142.33)	1,813	721	0.40	6
5	Gaari	Arnhem	QLD	(-11.167,132.92)	52	100	1.93	3
6	Jaako	Arnhem	NT	(-11.167,132.67)	777	721	0.93	4
7	Wurango	Arnhem	NT	(-11.25,132.083)	1,295	721	0.56	2
8	Iwaidja	Arnhem	NT	(-11.42,132.583)	259	321	1.24	5
9	Djalakuru	Arnhem	NT	(-11.5,133.057)	1,554	1,283	0.83	6
10	Oitbi	Arnhem	NT	(-11.5,132.5)	777	963	1.24	5
11	Tiwi	North	NT	(-11.5,130.83)	8,029	2,787	0.35	2
12	Mutjati	East Cape	QLD	(-11.583,142.75)	388	721	1.86	5
13	Unjadi	West Cape	QLD	(-11.58,142.58)	1,295	460	0.36	8
14	Nggamadi	West Cape	QLD	(-11.583,142.25)	1,942	690	0.36	3
15	Maung	Arnhem	NT	(-11.75,133.5)	1,295	585	0.45	4
16	Lotiga	West Cape	QLD	(-11.83,142.25)	1,036	271	0.26	5
17	Otati	East Cape	QLD	(-11.92,142.92)	777	1,143	1.47	4
18	Nango	Arnhem	NT	(-11.916,135.67)	777	721	0.93	2
19	Amarak	North	NT	(-11.917,132.83)	2,331	1,287	0.55	6
20	Ngathokudi	West Cape	QLD	(-12.08,142.417)	1,554	406	0.26	6
21	Tjongkandji	West Cape	QLD	(-12.08,141.917)	388	335	0.86	3
22	Diakui	Arnhem	NT	(-13,135.417)	5,698	1,900	0.33	3
23	Barara	Arnhem	NT	(-12.08,134.67)	518	721	1.39	3
24	Gambalang	Arnhem	NT	(-12.08,133.58)	1,554	716	0.46	3
25	Atjinuri	West Cape	QLD	(-12.17,142.67)	1,813	474	0.26	8
26	Tepiti	West Cape	QLD	(-12.17,142.17)	1,036	894	0.86	5
27	Nakara	Arnhem	NT	(-12.17,134.5)	518	721	1.39	4
28	Gunavidji	Arnhem	NT	(-12.17,134.17)	1,295	1,206	0.93	4
29	Ngardok	North	NT	(-12.17,132.5)	518	721	1.39	3
30	Jupangati	West Cape	QLD	(-12.33,141.83)	1,295	1,125	0.87	5
31	Dangu	Arnhem	NT	(-12.6,136.5)	2,590	782	0.30	3
32	Djangu	Arnhem	NT	(-12.3,136.5)	1,424	439	0.30	1
33	Djinang	Arnhem	NT	(-12.3,134.5)	518	259	0.50	6
34	Gadjalivia	Arnhem	NT	(-12.3,134.5)	518	259	0.50	6
35	Gunwinggu	Arnhem	NT	(-12.3,133.58)	7,252	2,746	0.38	10
36	Noreweilemil	North	NT	(-12.3,132)	1,036	540	0.52	3
37	Pakadji	East Cape	QLD	(-12.42,143.083)	3,367	1,582	0.476	5
38	Njuwathai	West Cape	QLD	(-12.42,142.3)	1,813	1,125	0.62	8
39	Ngormbur	North	NT	(-12.42,132.33)	2,072	1,198	0.58	6
40	Djerimanga	North	NT	(-12.42,131.42)	3,108	1,898	0.61	5

References

[1] Rao, A. R. *Smallpox* (Kothari Book Depot, Bombay, 1972).

[2] Hopkins, D. R. *The Greatest Killer: Smallpox in History* (University of Chicago press, 2002).

[3] Fenner, F., Henderson, D. A., Arita, I., Ježek, Z. & Ladnyi, I. D. *Smallpox and its Eradication* (World Health Organization, Geneva, 1988).

[4] Dixon, C. W. *Smallpox.* (J & A Churchill, London, 1962).

[5] Harper, G. J. Airborne micro-organisms: survival tests with four viruses. *Epidemiology & Infection* **59**, 479–486 (1961).

[6] Stallybrass, C. O. *The Principles of Epidemiology and the Process of Infection* (Routledge, London, 1931).

[7] Peterson, N. The natural and cultural areas of Aboriginal Australia: a preliminary analysis of population groupings with adaptive significance. *Tribes and Boundaries in Australia* 50–71 (1976).

[8] of National Mapping Australia, D. *Atlas of Australian Resources* (Department of National Development, Canberra, 1967).

[9] McCarthy, F. D. "Trade" in Aboriginal Australia, and "trade" relationships with Torres Strait, New Guinea and Malaya. *Oceania* **9**, 405–438 (1939).

[10] McCarthy, F. D. "Trade" in Aboriginal Australia, and "trade" relationships with Torres Strait, New Guinea and Malaya (continued). *Oceania* **10**, 80–104 (1939).

[11] Wurm, S. A. Linguistic research in Australia, New Guinea, and Oceania. *Linguistics* (1972).

[12] Kroeber, A. L. Basic and secondary patterns of social structure. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland* **68**, 299–309 (1938).

[13] Farris, W. W. *Population, Disease, and Land in Early Japan, 645–900*, vol. 24 (Brill, 2020).

[14] Crawford, M. H. *The Origins of Native Americans: Evidence from Anthropological Genetics* (Cambridge University Press, 2001).

[15] Dobyns, H. F. Estimating aboriginal american population: an appraisal of techniques with a new hemispheric estimate. In *The Atlantic Slave Trade*, 163–186 (Routledge, 2017).

[16] Watts, S. J. *Epidemics and History: Disease, Power, and Imperialism* (Yale University Press, 1999).

[17] Curson, P. *Times of Crisis: Epidemics in Sydney 1788-1900* (Sydney University Press, 1985).

[18] Gani, R. & Leach, S. Transmission potential of smallpox in contemporary populations. *Nature* **414**, 748–751 (2001).

[19] Arino, J. & van den Driessche, P. A multi-city epidemic model. *Mathematical Population Studies* **10**, 175–193 (2003).

[20] Stirling, E. C. Preliminary report on the discovery of native remains with an enquiry into a pandemic among Australian aboriginals. *Transactions and Proceedings of the Royal Society of South Australia* **35**, 4–46 (1911).

[21] Cleland, J. B. Some diseases peculiar to or of interest in Australia. *University of Sydney Medical Society Journal* **29**, 28–48 (1911).

[22] Cumpston, J. H. L. *The History of Small-pox in Australia, 1788-1908* (Albert J. Mullett, Melbourne, 1914).

[23] Cleland, J. B. Diseases among the Australian Aborigines. *Journal of Tropical Medicine and Hygiene* **31**, 54–56,67–70 (1928).

[24] Frost, A. *Botany Bay Mirages: Illusions of Australia's Convict Beginnings* (Melbourne University Press, 1994).

[25] Butlin, N. G. *Our Original Aggression: Aboriginal Populations of Southeastern Australia 1788-1850* (George Allen & Unwin, Sydney, 1983).

[26] Campbell, J. *Invisible Invaders: Smallpox and Other Diseases in Aboriginal Australia 1780-1880* (Melbourne Univ. Publishing, 2016).

[27] Tindale, N. B. *et al. Aboriginal Tribes of Australia: Their Terrain, Environmental Controls, Distribution, Limits, and Proper Names* (Australian National University Press, 1974).

[28] Bradshaw, C. J. A. *et al.* Stochastic models support rapid peopling of Late Pleistocene Sahul. *Nature Communications* **12**, 2440 (2021).

[29] Crabtree, S. A. *et al.* Landscape rules predict optimal superhighways for the first peopling of Sahul. *Nature Human Behaviour* **5**, 1303–1313 (2021).

[30] Williams, A. *et al.* Large size of the Australian Indigenous population prior to its massive decline following European invasion (2024). PREPRINT (Version 1) available at Research Square, doi:10.21203/rs.3.rs-5127915/v1.

[31] Radcliffe-Brown, A. R. Former numbers and distribution of the Australian Aborigines. *Offical Yearbook of the Commonwealth* **23**, 671–696 (1930).

[32] Birdsell, J. B. Some population problems involving Pleistocene man. In *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 22, 47–69 (Cold Spring Harbor Laboratory Press, 1957).

[33] Tindale, N. B. & Noone, H. V. V. Results of the Harvard-Adelaide Universities Anthropological Expedition, 1938-1939: distribution of Australian aboriginal tribes: a field survey. *Transactions of the Royal Society of South Australia* **64**, 140–231 (1940).

[34] Horton, D. *Aboriginal Australia [map]* (AUSLIG, 1999).