Epidemiology of the Plague of Athens

David M. Morens, M.D.
Robert J. Littman
University of Hawaii

...Sometimes...the barbarous rage
Of plague or pestilence, attends mans age,
Which neither force nor arts assuage;
Which cannot be avoided, or withstood,
But drowns, and over–runs with unexpected flood...
Draw back, draw back thy sword, O Fate!
Lest thou repent when 'tis too late,
Lest by thy making now so great a waste,
By spending all mankind upon one feast,
Thou starve thyself at last.

—From Thomas Sprat, "The Plague of Athens" (London 1676)

The cause of the celebrated epidemic in Athens (430–425 B. C.) is a subject of longstanding controversy. For nearly 500 years, scholars have tried unsuccessfully to identify the infection by comparing clinical features recorded by Thucydides to those of contemporary diseases.[1] At least 29 different disease theories have been advanced by hundreds of scholars. Adding to the controversy over its etiology was a 1985 publication by Langmuir *et al.* that revived an earlier influenza theory by explaining its incompatible clinical features as complications of toxic shock syndrome.[2] The national press popularized the controversy, including UPI wire service and *Newsweek* magazine coverage. Partially in reaction to this controversy, a panel was held at the annual meeting of the American Philological Association (Society for Ancient Medicine) where Langmuir presented his case, and various theories were debated without consensus. At that meeting co–author Morens outlined an

---

[1] Over 200 articles and books have been written in this time. Among the best current surveys are: J. Scarborough, "Thucydides, Greek Medicine, and the Plague at Athens: A Summary of Possibilities," *Episteme* 4 (1970) 77–90; J. C. F. Poole and A. J. Holladay, "Thucydides and the Plague of Athens," *CQ* 29 (1979) 282–300; and J. Longrigg, "The Great Plague of Athens," *History of Science* 18 (1980) 209–225. See most recently R. Sallares, *The Ecology of the Ancient Greek World* (London 1991) 221–294, esp. 243–266.

[2] A. D. Langmuir, T. D. Worthen, J. Solomon, C. G. Ray, E. Petersen, "The Thucydides Syndrome. A New Hypothesis for the Cause of the Plague of Athens," *New Engl. J. Med.* 313 (1985) 1027–30; L. Mercier, "Essai d' Interprétation de steriskomenoi de la 'Peste' d'Athènes," *Bull. Assoc. Guillaume Budé* 4th series 2 (1974) 223–6.

epidemiologic approach to identifying the cause of the Athenian epidemic. In this paper we suggest that epidemiologic aspects of the Athenian disease strictly limit possible modes of transmission and thus possible causes, ruling out influenza and many other previously suggested diseases. To assess the candidacy of the remaining diseases we suggest a systematic approach that includes (a) use of mathematical models (theoretical equations that purport to predict how epidemics will proceed under defined sets of conditions) and (b) comparison of the actual Athenian epidemic, of mathematically modelled epidemics of suspected diseases, and of epidemics of known diseases described and epidemiologically distinguished in well–characterized premodern populations. We argue that this process constitutes a conceptual framework from which proposed causes should be viewed, and that the results of this approach exclude most diseases suggested as having caused the Athenian epidemic.

Our method tries to identify, or at least limit, candidates for the Athenian disease by epidemiologic means in two successive stages: first we attempt to determine the mode of its transmission, and thereby to exclude diseases transmitted by other, inconsistent means; and secondly we attempt to find a "best fit" from the remaining candidate diseases, using (a) established mathematical models for respiratory disease transmission, and (b) data on premodern epidemics in characterized populations. As part of this process we examined Thucydides' text for aspects of descriptive epidemiologic data and synthesized these with information on such factors as climate, population density, and population growth dynamics, to construct a global picture of the epidemic according to the approach of modern epidemiologists conducting investigations of diseases of unknown cause. We compared the language and phraseology of Thucydides and Hippocrates, his contemporary, to look for equivocal or inappropriate medical terms or formulaic constructions, and we scrutinized Thucydides' clinical description for resulting ambiguities. We also evaluated scholarly publications concerning both the text and the meanings of Thucydides' description of the epidemic.[3] We considered textual matters, such as those bearing upon the likelihood that Thucydides could have observed what

---

[3]These included: H. Haeser, "Die Pest des Thukydides (430 v. Chr)," *Historische–pathologische Untersuchungen. Als Beiträge zur Geschichte der Volkskrankheiten,* vol I, Chapt. II (Dresden 1839) 32–57; J. Ehlert, *De Verborum Copia Thucydides Quaestiones Selectae* (Berlin 1910); W. Nestle, "Hippocratica" *Hermes* 73 (1938) 1–38; D. L. Page, "Thucydides' Description of the Great Plague at Athens," *CQ* 47 (1953) 97–119; A. W. Gomme, *HCT* II 145 ff.; C. Lichtenthaeler, *Thucydide et Hippocrate vus par un historien–médecin* (Geneva 1965); G. M. Parassoglou, unpublished diss. (Yale University 1968); A. Parry, "The Language of Thucydides' Description of the Plague," *BICS* 16 (1969) 106–118.

he described, sequelae that could have been rare, unrepresentative, or inaccurately described, the limitations of Thucydides' non–technical vocabulary in describing medical conditions, and the likelihood that concurrent conditions could have been misconstrued as being part of the symptomatology of the epidemic disease. In all cases we placed primary emphasis on epidemiologic information, drawing secondarily upon clinical information more liable to error. Using this framework, we attempted to identify the disease by identification of its mode of transmission, including evaluation of proposed causes of the epidemic using mathematical models of known candidate diseases, fitted with time/population data for Athens, and comparison of the epidemic behavior of candidate diseases to their behaviors in actual premodern epidemics. For this illustrative purpose we present the simplest of the mathematical models, that of Maia,[4] an early attempt to describe mathematically the so–called "Reed–Frost theory" of how epidemics behave in populations. It should be emphasized that our use of this particular mathematical model should not be construed as an endorsement of its validity. Definitive epidemic analysis using mathematical modelling is beyond the scope and intent of this paper. In fact, empirical validation of this and more sophisticated mathematical models is wanting for most diseases. Since each observed set of epidemic conditions is unique, validation for a particular disease in a particular epidemic setting would not necessarily allow generalization to the model's overall validity. Furthermore, the epidemic in Athens may have been quite different from epidemics of modern diseases. For example, high mortality may have affected the human contact rate in ways that cannot now be suspected. Newer more complicated mathematical models do a better job of taking into account such factors as the elements of chance and the changing contact rate. It is not clear, however, if these newer models are more successful in predicting and describing the course of actual epidemics of highly contagious diseases. The Maia model seems most appropriate for three reasons: (1) it is the best known of many models; (2) it is the simplest, and by far the easiest for non–mathematicians to understand, and (3) it requires fewer assumptions.

     Our approach emphasizes descriptive epidemiology, the characterization of population health events in terms of "person," "place," and "time," as described in standard epidemiology textbooks.[5] This method, based loosely on

---

[4]J. D. O. Costa Maia, "Some Mathematical Developments on the Epidemic Theory Formulated by Reed and Frost," *Human Biol.* 24 (1952) 167–200.

[5]For example, J. S. Mausner and S. Kramer, "Descriptive Epidemiology: Person, Place, and Time," *Epidemiology—An Introductory Text* (Philadelphia 1985) 118–153.

the approach of Hippocrates (*Airs, Waters, Places*), characterizes disease by its behavior in populations, rather than in individuals, as would be the case with a clinical approach. The epidemiologic approach includes determination of which types of individuals are susceptible to the disease (e.g. children vs. adults, males vs. females), where the disease occurs (e.g. in rural vs. urban locations, wet vs. dry areas), and when the disease occurs (e.g. winter vs. summer). Such an approach can characterize individual diseases, and in modern epidemiologic practice it is used to help identify the agents of epidemics of unknown cause. Since pre–modern clinical descriptions are by their very nature ambiguous, the epidemiologic picture may present a better means of identifying and understanding diseases from the past than dissection of their ostensible clinical features.

Epidemiologic Interpretation.

Person. Thucydides not only tersely states (2.51) that the disease "carried away all alike," but implies that it spared no population subgroup, including citizens, metics, slaves, and refugees. He does not mention clinical differences in the inbred citizenry, in which first cousin marriages were typical, nor occupational risk in food–handlers, hostlers, shepherds, or other persons with animal contact. The army and civilian populations were infected alike, although there may have been a somewhat lower case fatality rate in the cavalry than in the hoplite ranks. The cavalry presumably included the wealthier citizens. Thucydides says in one passage that persons with prior illnesses "all caught the plague in the end" (49). Alternatively, we cannot rule out the possibility that exposures associated with treatment of the prior diseases (e.g. exposures to physicians) were associated with an increased chance of acquiring the disease. Thucydides specifically links physicians to increased risk, without distinguishing between the risk of getting infected (attack rate) and the risk of dying after becoming ill (case fatality rate). Although Thucydides does not comment upon the Athenian diet during the sieges, deficiencies in total calories, protein, or most vitamins would have been unlikely because of grain warehouses and other storage capability in the city, and an unblockaded port (Piraeus) with large fleets of naval, commercial, and fishing vessels. But without fruits or vegetables inside the city walls, and with crops in the surrounding countryside either burned or appropriated by the occupying Spartan army, the Athenian diet would probably have been deficient in ascorbic acid (vitamin C). Vitamin A deficiency would be less likely unless some persons were also unable to obtain fish, fish products, fruits or vegetables.

Place. The origin of the epidemic cannot be determined with certainty. According to Thucydides (48.1) it was reputed to have originated in Ethiopia, spread to Egypt and Libya, and spread from there to the territory of the King of Persia [much of the Middle East]; before reaching Athens (47), "previous attacks...had been reported from many other places in the neighborhood of Lemnos and elsewhere." Thucydides states (47) that there was no record of any disease being so extensive or so destructive to men and that (54.5) Athens was hardest hit. Livy (4.20–21; 4.25.3–4;.4.30.8–10) describes epidemics in Rome in 433 and in 428. These may be part of the same pandemic. Plutarch (*Per.* 35.3) says the disease destroyed not only the Athenians, but also those who had dealings with their forces. Whatever the epidemic's origin, it was probably shipborne. Piraeus was the Mediterranean's major port at the time. Premodern epidemics and pandemics of many diseases (cholera, dengue, plague, smallpox) were typically spread by ships. Thucydides also says that the disease devastated Hagnon's naval expedition to Potidaea around July 430 B. C., at the height of the first epidemic wave. This expedition would have been under sail for about five days. After sailing, landing, marching and encamping, Hagnon's army suffered a paralyzing epidemic in Potidaea over a period of about six weeks. Many of the 3,000 Athenians already there subsequently became infected as well, although the besieged Potidaeans, reduced even to starvation and cannibalism, apparently did not. Thucydides also noted directionality of spread: after beginning in the port area of Piraeus, the disease spread up into the hills of Athens, separated from Piraeus by a narrow corridor between the Long Walls that was probably crowded with refugees from the countryside. Though it was in these refugee camps that the epidemic is said to have flourished, Thucydides did not distinguish between high attack rate (the epidemic incidence rate; i.e. the number of new cases per population at risk per time), high case–fatality rate (proportion of cases who died), or rapidity of epidemic progression. He linked the explosiveness in the camps, however, to "poor ventilation," consistent with either airborne spread, poor hygiene and sanitation, crowding, or any combination of these. His term "ventilation" probably says more about the (then) prevalent Greek belief in the miasmatic theory of disease causation than about an actual risk factor. Some authors have interpreted Thucydides to have meant that crowding was a risk factor for the disease in Athens. Gomme believes instead that Thucydides meant that once the epidemic broke out, it seemed all the worse because of the pre–existing miseries associated with over–crowding (*HCT* 158). Even so, Thucydides mentions crowding twice; here, and in his later comment about involvement of other densely populated towns (54). Potable water was supplied to Piraeus by

cisterns and to Athens by numerous wells spaced throughout the city. Athens also had cisterns, a river–fed aqueduct (from the Ilissus), and several springs, most of which were said to have had brackish water. The bedrock underlying Athens has porous limestone in the surface layer, underlaid by slate and marl. As noted above, grain storage was probable during the Spartan sieges, if not at other times. A curious item implicit in Thucydides' account of the war, and mentioned explicitly in one passage, is that despite frequent contact with the Spartan enemy, with the various allies in many regional cities, with the various expeditionary enemies, and with numerous allies, among the mainland cities of Greece only the Athenians were known to have suffered extensively from the devastating epidemic disease. Despite periodic encirclement of the Athenians by the Spartans, and frequent contact between the two enemies, the disease apparently either did not spread immediately beyond the city, or if it did, only when the Athenians themselves exported it. Thucydides (54.5) says that it did not enter the Peloponnesus to any extent and its "full force was felt at Athens, and, after Athens, in the most densely populated of the other places."

Time. The disease was not seasonal: it began soon after the Spartan army laid siege to Athens in late spring/early summer, probably in early May 430 B. C., a time of heat and humidity. Subsequent epidemic waves were noted in the summer of 428 B. C. and winter of 427–426 B. C. The epidemic apparently ended after 4 1/2 or 5 years, in winter 426–425 B. C. Just before the first epidemic wave, the Athenian population would have risen from at least 100,000 persons to about 300,000 or 400,000 persons as refugees streamed into the city. Gomme cites the base population of Athens as 155,000, composed of 60,000 citizens, 25,000 metics, and 70,000 slaves. Rostovtzeff believed that during the sieges the population rose to 315,000, and Major to over 400,000.[6] Many scholars appear to agree with the latter figure. The plague continued in Athens after its appearance in 430 B. C., even throughout the summer of 429 B. C., when Attica was not invaded at all. Neither was Athens under siege during the epidemic wave of winter 427–426 B. C., and the population would thus probably have been closer to 100,000, certainly less than 200,000, if refugees had returned to the countryside. Perhaps the most important set of clues Thucydides provides is that about the epidemic's duration: it continued uninterrupted for more than two, and perhaps for as many as four and a half to five full years (i.e. at the least from spring or early summer 430 B. C. to summer or fall 428 B. C.). Thucydides notes that it underwent an explosive

---

[6]A. W. Gomme, *The Population of Athens in the Fifth and Fourth Centuries B.C.* (Glasgow 1933) 22, 44; M. Rostovtzeff, *The Social and Economic History of the Hellenistic World* (Oxford 1941) I 95; R. H. Major, *Fatal Partners, War and Disease* (New York 1941) 9, 13.

rekindling in the winter of 427–426 B. C., continuing for a year after that. In reporting this latter epidemic wave Thucydides states "...the plague broke out among the Athenians for a second time. In fact, it had never entirely stopped...This second outbreak lasted for no less than a year, and the first outbreak had lasted for two years" (3.87.3). It is uncertain, then, whether at that point (winter, 427–426 B. C.) the epidemic had gone on continuously for three and a half full years (spring or summer 430 to winter 427/426 B. C.), or had actually stopped or had become barely detectable after two years of continuous activity (428 B. C.), only to have restarted again in the winter of 427/426 B. C., and then prevailed for another year thereafter. There can be little doubt, however, that it was prevalent for a very long time.

Other Epidemiologic Information. Thucydides clearly described a "virgin soil" epidemic. The novelty of the disease is supported by Thucydides' contention that it had not occurred previously, by signs and symptoms Thucydides claimed were distinctive, and by the rapidity with which it spread through the population. Population susceptibility is supported by the high attack rate, and an unvarying course in persons of different age, sex, and nationality, including an unsparing course in older persons. Although neither the incubation period, attack rate, or the case–fatality rate are known, in about July–August 430 B. C. the epidemic killed 26 per cent of an expedition of 4,000 hoplites in approximately 40 days (58). Elsewhere (3.87.3) Thucydides cites a final Athenian total of 4,400 hoplite deaths (34 per cent), and 300 cavalry deaths (30 per cent). This may be indirect evidence that the first epidemic wave (which killed 26 per cent of 4,000 hoplites) was by far the worst: assuming proportional mortality in the rest of the hoplites, 78% of the eventual 4,400 hoplite deaths would have occurred in the first wave. Thucydides did not say how many of them were ill and survived, and he would have had no way of knowing how many had been asymptomatically infected. In any case, the data suggest an attack rate between 25–100%, and a case–fatality rate more than 25% but less than 100%, since in the non–expeditionary population the disease was not invariably fatal. Thucydides himself survived, and there were apparently enough other survivors to suggest to Thucydides that survivors were never attacked twice. Thucydides makes several remarks that might seem to be clues to the mode of transmission, including comments that physicians were at increased risk, and that they became ill from tending the sick (the earliest surviving description of what has been cited as a concept of contagion). He also ambiguously implies the possibility of zoonotic (animal) disease: "...though there were many dead bodies lying about unburied, the birds and animals that eat human flesh either did not come near them or, if

they did taste the flesh, died of it afterwards. Evidence for this may be found in the fact that there was a complete disappearance of all birds of prey: they were not to be seen either round the bodies or anywhere else. But dogs, being domestic animals, provided the best opportunity of observing this effect of the plague" (2.50.1–2). Though Thucydides did not claim that anyone had seen ill dogs or birds, some authorities have cited this passage as proof of en– or epizootic involvement (low level circulation or outbreak transmission in animals). His empiric observation, however, is merely absence of birds and dogs about human corpses; he obviously expected to see them but did not. Furthermore, the structure of this passage seems to us to imply a perceived need to explain the unexpected absence of birds and dogs by offering a possible reason for it (that they became ill too, just as humans did), suggesting that the observation was not necessarily a fact Thucydides meant to preserve for posterity. It has even been suggested that Thucydides inserted this passage purely to bolster his contention that the epidemic disease was a new one; so new, presumably, that even the dogs and birds were too cautious to approach the carcasses.[7] Epizootic or enzootic disease is thus highly questionable.

The Text of Thucydides. Our analysis leads us to conclude that Thucydides' description of the epidemic disease is subject to potential error. A number of difficulties in interpreting Thucydides' text suggest caution in accepting his medical opinions, especially when they appear to conflict with factual or epidemiologic information presumably less liable to misinterpretation 2,500 years later. Thucydides' description of the epidemic seems to jumble physical signs and symptoms, epidemiologic observations, and historical facts. But it is his accuracy in describing signs, symptoms and clinical features that is the focal point of disputes about textual credibility. Translations of Greek lay terms into English medical terms for distinct signs and symptoms may be subject to considerable inaccuracy. Since there existed no standard or specific medical vocabulary to describe what he had observed, Thucydides would have applied "common" words to signs and symptoms, perhaps incompletely aware of the extent to which multiple meanings or ambiguities might later confound attempts to link the terms to medical concepts to be developed many centuries later, and perhaps even unaware of which common terms might have been favored by Hippocrates or other physicians.

     Thucydides may also have chosen words that had special or colloquial meanings since lost. Littré believed the account to have been written for the

---

7C. Anglada, *Étude sur les maladies eteintes et les maladies nouvelles pour servir à l'histoire des evolutions séculaires de la pathologie* (Paris 1869) 1–50.

"vulgar," and without scientific merit.[8] Harrison rues Thucydides' "[characteristic] over–condensation of style," which she believes "leads us to convict Thucydides of a real and unavoidable inexactness." She also notes Thucydides' penchant for inserting "vague addend[a]" to cover points about which he is uncertain.[9] Thucydides, she concludes, "always leaves perhaps rather much to the intelligence of his readers." Parry not only believed the account was hampered by a non–technical vocabulary, but claimed that Thucydides specifically avoided technical terms and wrote the epidemic account not for purposes of scientific accuracy, but to dramatize the tragic plight of the Athenians.[10] A "medical" term that scholars appear to consider least ambiguous is *phlyktainai*. It is often assumed that the epidemic disease was bullous because Thucydides used a word (*phlyktainai*) that had been used to describe spots on a loaf of bread, lesions on the hands of rowers, and other things. While this may be an excellent description of bullae, it might also be a description of other lesions for which the Greeks had no specific and unique medical word. Scholars have used a number of English translations of "phlyktainai": blains, blebs, blisters, bullae, eruptions, pimples, pustules, vesicles, and whelks. Hippocrates apparently used "phlyktainai" to describe not only bullae and vesicles, but pustules, burns, "sweat eruptions," and contact dermatitis, and others writers have suggested the term includes papules, plaques, and erythematous lesions as well.[11] Shrewsbury argues that while "phlyktainai" may typically mean "blisters," it is a general and inclusive term applied to any raised eruption of any sort.[12] A typical example can be seen in the translation of the Hippocratic word *sepedon* (not used by Thucydides). Some authors translate it as "gangrene", but Hippocrates used the word to describe cellulitis and purulent bacterial infection. We thus question with what assurance any such terms can be equated with modern medical terms, especially when certain theories of the cause of the epidemic hinge largely on their exact meanings. The theories suggesting smallpox, syphilis, and typhus, for example, are built around connection of such terms (e.g. *phlyktainai* in the case of smallpox) to modern terms for medical conditions.

---

[8]É. Littré, "Deuxième livre des épidémies," *Oevres complètes d'Hippocrate*, Vol. 5 (Paris 1839) 43–71.

[9]J. E. Harrison, "The Ancient City, Its Character and Limits" and "The Sanctuaries in the Citadel," Chpts. I and II in *Primitive Athens as Described by Thucydides* (Cambridge 1906) 5–36, 37–65.

[10]Parry (above, n. 3) 106–118.

[11]H. Keil, "The Louse in Greek Antiquity, with Comments on the Diagnosis of the Athenian Plague as Recorded by Thucydides," *Bull. Hist. Med.* 25 (1951) 305–23.

[12]J. F. D. Shrewsbury, "The Plague of Athens," *Bull. Hist. Med.* 24 (1950) 1–25.

It should be abundantly clear to any medical person reading Thucydides that the signs, symptoms, and features of the epidemic disease do not add up. Trying to put them together into a coherent picture of any disease, even a hypothetical extinct disease, leaves an impression not unlike trying to put together pieces of different puzzles: something, or perhaps some things, are clearly wrong. But which ones are they? From the medical perspective, how confident should we be in fitting together ostensible signs and symptoms when even Greek linguistic experts fail to agree on, and sometimes strongly disagree on, the meaning of important terms and phrases? What emphasis should be placed on Thucydides' signs and symptoms of the Athenian disease when we cannot decide between bleeding and congestion, pallor and jaundice, blindness and ocular damage, vesicles and other eruptions, bloody and watery diarrhea, convulsions and muscle spasms, retching and hiccuping, amnesia and dementia, lividity and flushing, colic and tenesmus, distal gangrene and some other loss of, or loss of the use of, the extremities? It is not even clear what killed the victims: Thucydides states in one passage (49.6) that an "internal fever" was the cause of death, or else diarrhea. But fever and diarrhea are non–specific symptoms of illness, not causes of death. Without knowledge of physiology, microbiology, pathology, or clinical medicine, what criteria would Thucydides have used for determining the cause of death? Could he have distinguished what people died *with* from what they died *of?* If not, his comments about deaths from either "internal fever" or diarrhea may be of no more relevance than to indicate either rapidly progressive courses, or more indolent courses ending with non–specific and agonal features like diarrhea. When scholars cannot even agree on whether Thucydides was referring, in one passage (49.6), to symptoms in the heart, or in the stomach, there may be a need for caution in any consideration of clinical features. Such problems may not be easily overcome by clinical or linguistic research efforts. For this reason, too, we favor the epidemiologic approach. Even so, we must be cautious in interpreting Thucydides' comments on the epidemiology of the disease. We doubt that Thucydides would have been able, for example, to distinguish such basic concepts as virulence and case fatality. We should not, therefore, over–interpret his words.

Moreover, some of the vocabulary problems that confuse us today may also have hindered Thucydides, though presumably to a lesser degree. Whatever trouble Hippocrates may have had with the lack of a precise medical vocabulary would have been considerably worse for Thucydides, who was not a physician, and thus would not have had experience observing and describing diseases. It should also be noted that despite some familiarity with the

Hippocratic emphasis on prognosis, and with certain Hippocratic terms, the vocabulary used by Thucydides to describe signs and symptoms of the Athenian epidemic differs from Hippocrates' descriptions of other diseases. An additional problem is that because Thucydides had the disease himself, his account of it may have been colored by subjective experiences. He might, for example, have stressed features he experienced, while de–emphasizing others he did not. Gomme suggests that while Thucydides took notes on the epidemic immediately after his own illness, he fleshed out the description of the epidemic much later. If so, did he dress up the account with Hippocratic trimmings? It is also clear that an implicit aim of his brief epidemic account was to gain credibility with physicians by presenting the disease description in a way that appealed to contemporary medical thought, and that fit (then) current medical theories. His assertion that during the epidemic patients succumbed on the seventh or ninth day, for example, is taken verbatim from Hippocrates, who espoused the view that certain conditions led to demise on those specific days (e.g. *Aphorisms* IV 36; IV 64). Or does the Hippocratic statement (*Aphorisms* III 21) on the effect of time of the year on disease, "in summer...we must also expect...gangrene of the genitalia," influence Thucydides' account of loss (of use) of the genitalia due to the disease? Although Thucydides claimed to be most interested in diagnosis, he tailored at least a part of his description to the prognosis–oriented expectations of the Hippocratic school, this despite the fact that he and his contemporaries were well aware of the prognosis. Thucydides' tendency to describe those disease features of greatest interest to Hippocratic followers, including formulaic phrases interspersed in the text, casts some doubt on the validity of the individual signs, symptoms, and features mentioned.

Finally, the organization of Thucydides' text on the Athenian epidemic suggests an unsatisfactory synthesis of multiple aims, and leaves the impression, if examined carefully, that Thucydides was not at all sure what the features of the disease were, or how to select, present and organize them in narrative form. After beginning his discussion with several paragraphs of descriptive epidemiology, he next describes the onset of illness, and then an apparent chronological course of disease progression: "..It began with...The next symptoms were...and before long...next the stomach..." (49). This chronology is then abruptly interrupted by descriptions of apparently different courses taken by the illness: "Most patients suffered...Whenever it settled...In some cases...in others...many...actually [jumped into wells], the majority...or else...In most cases there were...this sometimes ended...but sometimes continued..." (49). Later, Thucydides appears to be describing complications and sequelae,

though he would not have had such concepts in mind. He states, for example, that "...if people survived this critical period, the disease descended...some, too, went blind...There were some also who...suffered from a total loss of memory..." (49). The description obviously lacks any means of distinguishing between different courses of the disease, complications, sequelae, and supervening infections, had any or all of these occurred.

Thucydides also contradicts at least some of his earlier descriptions by introducing a "head–to–toe" organizational structure that competes with his initial chronology: "...first settling in the head, went on to affect every part of the body in turn..." (49). In fact, the competing concepts of chronology and head–to–toe progression are interwoven unsatisfactorily throughout his description. In our view, the disease features seem difficult or impossible to fit. Excluding consideration of atypical courses and complications, sequelae, and supervening illnesses, it is probably fair to generalize in saying, for example, that potentially fatal gastrointestinal diseases (e.g. cholera, shigella dysentery) do not cause the sorts of respiratory symptoms Thucydides describes and, on the other hand, that fatal respiratory diseases (e.g. influenza) do not cause the types of gastrointestinal features Thucydides records. In a similar vein, diseases that make people sneeze normally do not make their genitalia fall off. Previous critics, grappling with such paradoxes, have sought one of three escapes: (1) either "the" disease was many different diseases confused as one, or (2) it was just as Thucydides described it, but is now extinct, or (3) it was one disease after all, but Thucydides erred in his description of one or more of the features. We believe the last of these possibilities. On genetic and evolutionary grounds, we strongly doubt that the disease could have become extinct. While most microorganisms evolve quickly, they are highly adapted to humans or other hosts that evolve slowly, thus avoiding the threat of extinction unless the hosts become extinct themselves. Measles, for example, has apparently been prevalent for at least 1,000 years, plague for at least 2,000 years, and poliovirus diseases for at least 3,000 years. Each of these diseases appears not to have changed at all. Even smallpox has only undergone one recognized change in clinical form in its 3,000 or more years of prevalence, and this change (a reduction in virulence) was predictably in the direction of assuring its survival, rather than hastening its extinction. There is thus little empiric or theoretical reason to suspect extinction of any human disease except by purposeful eradication, as with smallpox. Although it is possible that other concurrent diseases contributed to some of the supposed signs and symptoms, we also doubt two or more diseases because of the difficulties in explaining a remarkably concordant time course, because Thucydides himself and the

Athenians in general apparently believed it was one disease, because it was confined to the Athenians, and because its reappearance over a five year period seemed to be associated with the same features.

From the above information we may make general inferences about the mode of transmission of the Athenian epidemic, eliminate most candidate diseases from consideration on this basis, and then make additional inferences about the suitability of diseases that remain. We compared three possible modes of transmission to the observations of Thucydides: common source, person–to–person (enteric, inoculation, and aerosol/respiratory), and "reservoir–associated". The term "reservoir" is taken to mean an animal, insect, or environmental source in which an infectious agent is maintained when not infecting humans.

Common–Source Acquisition. Common–source acquisition is not consistent with epidemic persistence over several years time, or Athenian exportation to Potidaea. In open populations, common source epidemics are nearly always foodborne or waterborne. Ergotism has been suggested as a cause of the epidemic,[13] in part because it causes peripheral gangrene, but there is no reasonable explanation for such massive and simultaneous contamination of all grain sources, more frequent (or worse) illness in refugees, lack of differences in illness by age, and documentation of geographic spread within the city. Grains would probably have been consumed in less than two to five years, especially during a siege characterized by unimpeded access of the Athenian fleet to the many Aegean grain–producing areas. Contamination of harvests in so many successive years, from such varying sources, is difficult to imagine. The same grains were also presumably consumed by others who did not experience epidemic disease, including the Spartans, who had appropriated the Athenian crops. Similar arguments rule out other foods in the Greek diet, including other grains, fish, and oils. As noted above, destruction or appropriation of ascorbic acid sources would probably have placed the Athenians in a state of vitamin C deficiency, at least temporarily. But the epidemic began too soon after the probable cut–off of ascorbic acid–containing fruits and vegetables to implicate scurvy. It first broke out at the beginning of the Spartan siege (late spring or early summer 430 B. C.), making scurvy an unlikely cause if, as we assume, it would take at least several weeks, if not months, for symptoms to develop. Scurvy could have altered or confused the

---

[13]A.–P. Read, "Épidémies occasionées par l'usage du seigle ergoté," *Traité du seigle ergoté dans lequel on examine les causes de cette excroissance végétale, les moyens de la prévenir, les résultats de l'analyse de ces grains, leurs effets sur les animaux, les maladies epidémiques qu'occasione leur usage, & le traitement qu'elles exigent*, Part 3 (Strasbourg 1771) 52–93.

clinical picture later on, however. Symptoms of scurvy, documented since biblical times,[14] are similar to some of those described by Thucydides (e.g. fever, oral bleeding, red eyes, hyperkeratotic skin lesions, petechiae, impaired wound healing, melena, and irritability/agitation) and could have produced features mistaken as disease complications. Recently, clinically inapparent vitamin A deficiency has been associated with worsening of severity of measles and other infectious diseases.[15] But despite its association with blindness, skin lesions, and diarrhea, vitamin A deficiency seems unlikely because of the probable presence in Athens of vitamin A–containing products, including fish, during the epidemic.

Waterborne disease is even more difficult to defend, since the epidemic began in Piraeus, a port whose cisterns were unconnected to other water sources and unassociated with sewage channels, thus providing no opportunity for wholesale contamination of potable water. Furthermore, the disease spread from Piraeus to Athens. With most of the Athenian wells on high ground, and with a soil underlaid with porous limestone, the direction of groundwater flow in 430 B. C. would have been, as now, downhill: any contaminant introduced into the water would have flowed downhill towards Piraeus (or to the north and west), not uphill towards Athens. There would also seem to be no practical way (except, conceivably, by avian contamination) for the many hundreds of Athenian wells to be simultaneously cross–contaminated at the onset of the epidemic, or to remain contaminated over such a long period of time. Historically, massive enteric disease epidemics have usually been associated with sophisticated water and sewage systems, e.g. the London cholera epidemics of the 1840s and 1850s. Zoonotic diseases that are also associated with "common source" exposures (e.g. anthrax) are considered below under the heading of "Reservoir Diseases." Common source acquisition of a single epidemic disease in 430–425 B. C. can otherwise be ruled out, and with it etiologic theories for cholera, dysentery, ergotism, shigellosis, scurvy, and typhoid fever.[16]

---

[14]J. H. Swanson, "Evidence of Scurvy Among Ancient Hebrews" *Bull. Hist. Med.* 15 (1944) 352–8.

[15]N. S. Scrimshaw, C. E. Taylor, J. E. Gordon, *Interactions of Nutrition and Infection* WHO Monograph Series. No. 57 (Geneva 1968); G. D. Hussey and M. Klein, "A Randomized, Controlled Trial of Vitamin A in Children with Severe Measles," *New Engl. J. Med.* 323 (1990) 160–4; L. Rahmathullah, B. A. Underwood, R. D. Thulasiraj, *et al*, "Reduced Mortality Among Children in Southern India Receiving a Small Weekly Dose of Vitamin A," *New Engl. J. Med.* 323 (1990) 929–35.

[16]Cholera: F.–R. Chateaubriand, *Mémoires d'outre–tombe* (Paris 1849). Dysentery: F. Kanngiesser–Neuenburg, "Die seuche des Thukydides (Typhus exanthematicus)," *Münchener Med. Wochenschrift* 63 (1916) 1627. Shigellosis: C. G. Ray, "Death in Athens: The epidemic

Person–To–Person Transmission. Of the three basic subcategories of person–to–person transmission, enteric ("fecal–oral"), inoculation (venereal and non–venereal), and respiratory/aerosol, the first two can be ruled out. Enteric transmission is even less likely than common–source acquisition of enteric organisms, already discounted. Though organisms transmitted by the waterborne route are typically also transmitted enterically (e.g. cholera), in open populations enteric transmission is, for obvious reasons, much less efficient in bringing microorganisms in rapid contact with large numbers of people. A purely enteric disease (i.e. one not simultaneously waterborne) could not possibly spread through tens of thousands of people in a few weeks. The same can be said of venereal inoculation, even assuming multiple introductions from returning navies. Thucydides' alleged description of "gangrene" of the genitalia may, as noted above, actually refer to some other condition. Non–venereal transmission by inoculation is also unlikely. An inoculum source widespread enough to cause such an explosive epidemic in over 100,000 people is difficult to imagine. Most human non–venereal inoculation diseases are zoonotic, as discussed below. Exclusion of person–to–person enteric and inoculation diseases eliminates, as noted above, cholera, "dysentery," shigellosis, and typhoid fever, and also eliminates syphilis, proposed as a cause of the Athenian epidemic in the mid 19th century.[17]

Consistent with person–to–person aerosol/respiratory transmission is the explosiveness of the Athenian epidemic (with high attack rate and rapid spread); the apparent lack of correlation of illness and either age, sex, or occupation; an association with crowding; and the supposed predominance of upper respiratory symptoms at illness onset. But most respiratory diseases, including virtually all that occur in explosive epidemic form, rapidly die out in closed crowded populations. Universal susceptibility may initially lead to explosive epidemic progression, but as the epidemic continues, the declining availability of susceptible persons slows further progression, until the epidemic ends for lack of new susceptibles: depending upon the ease of disease transmissibility either all persons have become infected and have, therefore, died or become permanently immune, or else only some persons have become

of 430 B.C.," Symposium 2 March 1984, Arizona Health Sciences Center. Tape A/7 114, Media Department, Health Sciences Center Library, Tucson, Arizona. Scurvy: E. H. Smith, "The Plague of Athens" in S. L. Mitchell, E. Miller, and E. H. Smith, *The Medical Repository*, 2nd ed., Vol. I (New York 1800) 3–32. Typhoid fever: C. Pfeufer, "Geschichliche Bedeutung des Scharlachs," *Der Scharlach. Sein Wesen und seine Behandlung, mit besonderer Beruchsichtigung des 1818 zu Bamberg herrschenden Scharlachs* (Bamberg 1819) 1–10.

[17]J. Rosenbaum, *Geschichte der Lustseuche im Alterthume* (Halle 1845).

infected, and the rest are protected by "herd immunity," a population immunity threshold, characteristic of each disease, that blunts or prevents epidemic transmission. Respiratory disease epidemics in distinct and crowded populations thus normally peak and end quickly, progressing with predictable, even mathematical regularity. But despite a closed population with both limited immigration and numerical attrition due to war, Thucydides stated unequivocally that the Athenian epidemic went on for at least two, and probably four or five years. Furthermore, despite considerable contact between the two armies, the Spartans appear not to have been affected by the disease. To invoke a respiratory epidemic of explosive onset, it is necessary to explain both its persistence over long periods of time and its apparently imperfect transmissibility to other persons with whom the Athenians undoubtedly came into contact during the epidemic years.

To understand better the behavior of respiratory diseases in human populations, it is helpful to consider certain epidemiologic concepts that bear upon the description of the Athenian epidemic. Since there is variation in both the incubation periods of diseases, and the percentages of the population susceptible to diseases at any given time in their epidemic courses, the rate of disease spread varies also. Incubation periods often tend to approximate the time from infection until the infected person becomes contagious, which may in turn approximate the serial generation times—the mean time intervals between peaks of successive epidemic waves as observed in actual epidemics. If a disease takes two days to render its victim capable of transmitting to others, it may spread faster than a similarly infectious disease that takes 10 days to become contagious. But if, for example, the two day disease infects only 10% of those exposed to it, while the 10 day disease infects 20%, their rates of progression in a population of susceptible persons may be difficult to predict without mathematical formulas. Using these so–called "mathematical models" we can predict for any contagious disease mathematical patterns of occurrence over specified time periods under given conditions (e.g. population size and degree of crowding). We can apply such theoretical patterns of specific diseases to the Athenian epidemic, as described by Thucydides, to look for concordance.

To evaluate further respiratory transmission we applied simple versions of "mathematical models" of known diseases, using information on Athenian population size, crowding, and susceptibility, to predict the time it would take for them to die out. Respiratory diseases of short incubation (influenza) and long incubation (measles, smallpox) are compared below, using the

straightforward deterministic mathematical model of Maia. The equation used to generate the predicted epidemic curves is:

$$C_{t+1} = S_t (1 - q^{C_t})$$

        where:

C_t =   number of cases of the disease at time t, the beginning of the epidemic;

C_{t+1} =   number of cases of the disease at time t + 1

t   =   the chosen time interval; here the serial generation time;

S_t =   the number of susceptible persons at time t

q   = $1 - p$

p   =   the probability of 'adequate contact' (i.e., adequate to cause infection) between any two individuals per time interval t.

For each disease the following assumptions are made: (1) there is universal susceptibility at the outset, with infection either killing or conferring permanent immunity (assumptions about case fatality thus have no bearing on the epidemic curves, since death and long–lasting immunity are equivalent barriers to further transmission); (2) the serial generation time is taken to be 4.5 days for influenza, 12 days for smallpox, 14 days for measles, and 19 days for streptococcal infections such as scarlet fever;[18] (3) the total population figure for Athens is taken to be the minimum of 100,000 as noted above (during most of the early war years the population probably alternated between about 100,000–200,000 and 300,000–400,000 as refugees entered and left the city during the siege[19]); and (4) a range of "adequate contact" numbers is selected to approximate conservative estimates of the frequency of interpersonal contact under conditions of known severe crowding. Although an adequate contact figure is difficult to arrive at with precision for any situation, we can estimate it (only very roughly) from at least two sources: data on population crowding indices, and from empiric back–calculation using Thucydides' information about the outbreak in Hagnon's naval expedition.

The area of Athens/Piraeus in 430 B. C., including the sparsely populated port and storage areas, was about four square miles, suggesting a population density ranging from 25,000 persons per square mile (about the same as New York City today), to about 100,000 per square mile (like modern Delhi). These modern comparison figures, of course, describe populations living in high–

---

[18]R. M. Anderson, "Directly Transmitted Viral and Bacterial Infections of Man," *The Population Dynamics of Infectious Diseases: Theory and Applications* (New York 1982) 1–37.

[19]It has been argued, however, that after the first siege large numbers of persons from the countryside probably remained in the city year–round (Gomme, *HCT*).

and low–rise buildings, which greatly reduce the degree of crowding. It is also clear that Athenians were not equally dispersed, leading to pockets of even greater crowding. Webster estimates that 30% of the square mile of Athens proper was normally unoccupied.[20] Thucydides implies that the narrow passage between the Long Walls connecting Athens and Piraeus (less than a square mile) was heavily populated by the refugees, and we may speculate that Piraeus (about two square miles) was less populated. Even if the minimum estimated population of 100,000 had been maximally dispersed over the entire land mass of Athens/Piraeus, each individual would only have occupied a space equivalent to a square patch of ground considerably less than 35 by 35 feet in size. Assuming a more probable population of 400,000 would assign each person a square patch of only 16 feet on a side, about the length of a compact automobile. But in reality, population dispersion is never maximal, so that considerably greater crowding than these estimates would be expected. For example, in crowded populations persons are crammed into houses, rooms, and public places, while other areas are unoccupied, leaving clusters of densely crowded areas.[21] Indeed, if Xenophon is correct in assuming 10,000 houses in Athens, there may have been at least 10 persons per average household before, or as many as 40 during the sieges, the majority of whom were presumably crowded into the refugee camps. (Estimates of 400,000 Athenians in 10,000 dwellings have been widely accepted.)[22] Attica itself, or at least the part surrounding Athens from which the refugees came, is only thought to have contained about 250 square miles of land. Another potential problem in interpretation is that in Athens/Piraeus the epidemic would probably not have been detected until some time after it actually began, confounding interpretation of Thucydides' comments on its duration. However, there is no such difficulty with the Hagnon expeditionary data if, as assumed here, the expedition would not have sailed from Piraeus with hoplites suffering from the disease on board. Thus, in estimating adequate contact by this means, no assumptions need be made about late epidemic recognition. In 4,000 expeditionary hoplites, plus 3,000 hoplites previously encamped in Potidaea

---

[20]T. B. L. Webster, "Attica and Its Population," *Athenian Culture and Society* (Berkeley 1973) 35–57.

[21]Cf. Webster 40: "The city of Athens was...a jumble of narrow streets with houses, private or partly industrial, and shrines closely huddled together...in such a crowded city news spread fast, and one could not help knowing what one's neighbor was doing...The very crowding of Athens...made meeting certain."

[22]R. E. Siegel, "Epidemics and Infectious Diseases at the Time of Hippocrates. Their Relation to Modern Accounts," *Gesnerus* 17 (1960) 77–98.

(Phormio and his 1,600 men had already left), the epidemic curve went from the index case to extinction in about six weeks.

In the Maia equation the calculated value of 'p', which changes as the epidemic proceeds (i.e. as 't' increases), is a function of two independent probabilities: the probability that any two individuals, regardless of immunity status, will contact each other during a serial generation time, and the probability of disease transmission should contact between an infectious person and a susceptible person occur. If, for example, during a smallpox epidemic an average infectious Athenian came "in range" of 25 persons per 12 days, and was capable of transmitting the infection to only 20% of those who were susceptible, there would be five "adequate contacts" per 25 susceptible persons infected every 12 days. Because estimation of the chance of adequate contact is somewhat speculative, in the illustrations below we present a range of adequate contacts (20, 10, 5, and 2) that we consider to be characteristic of low to extremely low transmissibility. A rough independent check on the reasonableness of these estimates may be found in the form of a "back–computed" adequate contact number from information Thucydides provided on Hagnon's naval expedition to Potidaea. However, such an example is illustrative rather than definitive: we have no reliable information concerning different epidemic conditions in Athens versus the Potidaean camps. A mathematical description of a single occurrence may be inadequate to predict what would happen in similar circumstances or with repeated trials. In Hagnon's expeditionary outbreak, Thucydides reported fatal disease in 1,050 (26%) of a 4,000 man hoplite force in about six weeks. Application of the Maia model to this outbreak, back–calculated with a 12 day serial generation time, would suggest that an adequate contact number of at least 30–40 would have been required to end the epidemic in six weeks, even assuming that (a) none of the hoplites had yet become immune two months or so into the epidemic (Figure 1), and that (b) the 3,000 previously encamped hoplites were not involved. If they were involved, as seems almost certain, the adequate contact number would have had to have been even greater for the epidemic in all 7,000 men to extinguish in six weeks. Thus the essential conservatism of our range of estimates of adequate contact numbers appears to be confirmed. Note that in our calculations an adequate contact number of 10 for a disease with a 12 day serial generation time (like smallpox) is equivalent to an adequate contact number of 4 for a disease with a 4 1/2 day serial generation time (like influenza). That is, we assume constant rates of contact between people.

The results of mathematical modelling in predicting the duration of epidemics of various transmissible diseases proposed as causes of the Athenian

epidemic should not be surprising to scholars who have studied actual epidemics in defined premodern populations. For example, under virtually any condition of crowding and contact believed to have prevailed in Athens of 430 B. C., influenza would have spread through the entire population and then died out in about nine weeks (Figure 2). Similarly, epidemic smallpox is predicted to have extinguished in but a few months (Figures 2 and 3). With 10 adequate contacts, for example, epidemic smallpox is estimated to last less than four months, leaving only about five persons untouched and still susceptible (Figure 3). With five adequate contacts, smallpox would be expected to die out in about five months, leaving only 698 persons still untouched (Figure 3). This is obviously too few to support subsequent epidemic transmission, even with a high birthrate in Athens. (The birthrate in ancient Athens is unknown; but even in the poorest countries today it normally ranges from only 3–5% of the entire population per year). Since the disease was not solely a pediatric disease in its later years, however, a large number of adult susceptibles must have existed later on. Even had smallpox in Athens been a disease of extraordinarily low transmissibility (two adequate contacts), the mathematical model does not predict disease persistence in Athens for as long as a year, let alone two to five years (Figure 3), though in this scenario 20,318 persons would remain susceptible at the end of the epidemic. Measles, with a slightly longer serial generation time (14 days) would likewise have rapidly been extinguished in Athens unless it was of extremely low transmissibility (Figure 2). This is in accord with the accepted dogma that year–round measles transmission does not occur in populations under about 300,000–500,000 persons, even with modern sanitation and lack of crowding.[23] Epidemic streptococcal disease (scarlet fever, erysipelas,[24] or both) is more difficult to exclude because of the lack of information about its behavior in susceptible populations.[25] By the time scarlet fever was distinguished from diphtheria (after Bretonneau's 1826 description

---

[23]J. A. Yorke, N. Nathanson, G. Pianigiami, J. Martin, "Seasonality and the Requirements for Perpetuation and Eradication of Viruses," *Am. J. Epidemiol.* 109 (1979) 103–23.

[24]Scarlet fever: F. Schnurrer, *Die Krankheiten des Menschengeschlects historisch und geographisch betrachtet. Chronik der Seuchen, in Verbindung mit den gleichzeitigen Vorgängen in der physische Welt un in der Geschichte der Menschen* (Tübingen 1823). Erysipelas: J. Malfatti, "Beschreibung eines bösartigen Scharlachfiebers, welches zu Wien im Jahre 1799 unter den Kindbetterinnen geherrscht hat, nebst einigen Bemerkungen," *Journal der practischen Arzneykunde und Wundarzneykunst* 4 (1801) 120–152.

[25]A. R. Katz and D. M. Morens, "Severe Streptococcal Infections in Historical Perspective," *Clin. Infect. Dis.* 14 (1992) 298–307.

of diphtheria,[26] and more universally after the 1854–1859 diphtheria pandemic), it had already become endemic in urban centers of most developed nations. But with a prolonged mean serial generation time (19 days) due in part to long–term infectivity (despite its short incubation period), streptococcal disease is perhaps, on epidemiologic grounds, a candidate for further research, despite apparent clinical inconsistencies. The mathematical model suggests, however, that in part because of its long mean serial generation time, epidemic streptococcal disease could have persisted in Athens longer than measles or smallpox. Assuming, for example, 10 adequate contacts per 12 days (equivalent to 16 contacts during the 19 day serial generation time of scarlet fever), the epidemic would have died out in less than five months (Figure 2). But streptococcal disease is one of the few infectious diseases that can attack persons more than once. This would facilitate disease persistence for several years. However, against identification of this disease is Thucydides' claim that the disease never attacked twice.

This mathematical model of diseases transmitted solely by the respiratory route does not adequately explain the Athenian epidemic. Even if they had once been, for some obscure reason, poorly transmissible, none of these diseases (e.g. influenza, measles, smallpox) would have persisted two to five years, as did the Athenian epidemic. Etiologic theories of diseases transmitted purely by the respiratory route (including some with combined inoculation transmission) are thus difficult to accept as causes of the Athenian epidemic, among them influenza and influenza–associated diseases (including Guillain–Barré syndrome and influenza complicated by staphylococcal toxic shock syndrome), measles, meningitis, smallpox, staphylococcal diseases (including toxic shock syndrome), and probably the unidentified "sweating sickness" of 16th century England and continental Europe.[27] The only windows of possibility for respiratory diseases are provided by either unusual mechanisms for extra–human persistence, or parallel explosive urban/non–explosive rural transmission. Both of these possibilities are addressed below.

---

[26]P.–F. Bretonneau, *Des inflammations spéciales du tissu muqueux, et en particulier de la diphthérite, ou inflammation pelliculaire, connue sous le nom de croup, d'angine maligne, d'angine gangréneuse, etc.* (Paris 1826).

[27]Meningitis: J. L. Lefèvre–Douvillé, *Essais Médico–littéraires sur les Anciens* (Paris 1858). Smallpox: A. Kircher, *Scrutinium physicomedicum contagiosae luis quae dicitur pestis* (Rome 1658); J. Freind, "To Dr. Richard Mead," *Two Epistles to Dr. Richard Mead: The One Concerning Purging in the Second Fever of the Confluent Small–pox: The Other Concerning Some Particular Kinds of the Small–pox* (London 1729) 1–120; R. J. Littman and M. L. Littman, "The Athenian Plague: Smallpox," *TAPA* 100 (1969) 261–75. Sallares (above, n. 1) 244–262. Sweating sickness: J. Caius, *A Boke, or Counseill Against the Disease Commonly Called The Sweate, or Sweatyng Sicknesse* (London 1552).

Reservoir–Associated Acquisition. The fact that the Athenian epidemic began explosively and lingered and resurfaced in a closed population over the course of two to five years without disappearing, and perhaps without greatly affecting others outside the city, suggests an extra–human mechanism for persistence of the agent responsible for the disease. Continuous infectious or non–infectious common source exposures (e.g. a waterborne enteric agent, or a foodborne source such as ergotism) have already been discounted, leaving infectious diseases with either animal reservoirs or insect vector reservoirs, or both. Many such diseases in which the causative organisms have established zoonotic cycles involving man as an accidental or "dead end" host (e.g. glanders and rabies), are not otherwise consistent with the epidemic. But for those in which the organism is co–adapted to humans or other primates (e.g. yellow fever, dengue, Rift Valley fever), persistence in the vector host for prolonged periods with or without epidemic or endemic activity is typical. Of the latter diseases, many produce remarkably explosive epidemic curves, surpassing in explosivity even those of influenza. Several of these reservoir diseases, addressed below as possible causes of the Athenian epidemic, meet two of what we consider Thucydides' most basic criteria: long–term persistence and explosive/re–emergent potential. "Reservoir–associated" disease transmission appears to be the only "pure" mode of transmission consistent with the observations of Thucydides. The principal argument against it is Thucydides' implication of person–to–person transmission, including his note that the epidemic raged in the most crowded areas. However, as discussed more fully below, vectorborne diseases frequently mimic the epidemiology of explosive respiratory diseases to a remarkable degree, even in their marked association with crowding.

Discussion.

Epidemiologic analysis of the Athenian epidemic is consistent with an infectious agent associated with either (a) an animal or insect reservoir, or (b) respiratory transmission combined with a "reservoir–like" mechanism of persistence. This conclusion is strongly supported by Thucydides' documentation that the epidemic continued without interruption for two to five years in Athens without noticeably affecting either the surrounding Spartan army or without, presumably, affecting many or most of the hundreds of thousands of other persons the Athenians would have come into contact with during the war years, and by the fact that when the Athenians did export the disease outside the walls of Athens (e.g. to Potidaea) only they appear to have been affected by it in any great degree. These two possible categories of

transmission of the agent of the Athenian epidemic are discussed in greater detail below.

Respiratory Transmission. As noted above, diseases transmitted only by the respiratory route can be discounted since they do not "over–winter," and would have spread rapidly in the extraordinarily crowded Athenian population. Comment on several of the specific disease examples follows. Despite our caution in interpreting mathematically derived data, it is worthwhile to note that in modern times the validity of certain mathematical models in predicting the course of influenza epidemics has been strongly supported by observational and empirically derived data.[28] There is thus reason to expect that the course of such an epidemic in a natural population, even one occurring more than 2,000 years ago, would be much as the models predict. Rapid extinction in Athens is also supported by consideration of hundreds of influenza epidemics documented over the last three centuries, including those reported by countries, states, counties, and cities. In 1918, for example, influenza spread through, and died out in Newark, New Jersey (population 435,000) in about 12 weeks. A similar pattern was observed in most American cities of this size in 1918. The 1918 epidemic, the most highly fatal ever recorded, spread through and died out in the entire United States—with a land mass nearly a million times larger than that of Athens in 430 B. C., with 10,000 times the population, and with a crowding index only one two thousandth as much—in about six months. The Athenian population would probably have had lower standards of sanitation and hygiene, no awareness of barrier protection or isolation, lower standards of medical care, and a greater burden of concurrent diseases and instances of under–nutrition. With respect to the recent theory that influenza complicated by a toxic shock syndrome–like disease caused the Athenian epidemic it might also be added that explosive staphylococcal epidemics are not known to occur. Only about 10–20% of persons can be chronically colonized with staphylococci; in open populations this apparently proceeds over prolonged periods. Colonization with some unique strain, e.g. a toxic shock syndrome toxin–producing strain, would be even less rapid were it to occur. Invoking influenza/toxic shock syndrome as the cause of the Athenian disease requires three epidemics: of influenza, of toxin–producing staphylococci, and of impetigo. A basic tenet of epidemiology is that unknown epidemics are rarely caused by two diseases. The lack of association of impetigo with toxic

---

[28]N. T. J. Bailey, *Mathematical Theory of Infectious Diseases and Its Applications* (London 1975); C. C. Spicer, "The Mathematical Modelling of Influenza Epidemics" *Br. Med. Bull.* 35 (1979) 23–28.

shock syndrome and the lack of association of the Athenian epidemic with pneumonia are additional factors against this possibility.

Although measles is more difficult to exclude on epidemiologic grounds, mathematical models suggest that it, too, would not have persisted for two to five years in Athens. This conclusion is also supported by contemporary empiric observations and accepted mathematical modelling studies; even in relatively uncrowded populations 300,000–500,000 persons are considered necessary to sustain measles transmission. As shown in Figure 2, unless it was less transmissible in a virgin population in 430 B. C. than is observed today, measles would have extinguished in a matter of months. When in addition we consider the crowding and sanitary conditions cited above, fitting measles and the Athenian epidemic becomes increasingly difficult.

Smallpox is the most difficult of the respiratory diseases to exclude, not the least because it has a long incubation period and is less transmissible than either measles or influenza. In considering smallpox in our mathematical model, we assumed a 12 day serial generation time, universal susceptibility, a range of 2–20 adequate contacts, and a minimal base population/crowding index of 100,000 persons per four square miles. Under such assumptions the mathematical model indicates that a smallpox epidemic would die out in less than 11 weeks given 20 adequate contacts; in 22 weeks given five adequate contacts; or, at the theoretical extreme, as long as about 11 months given two adequate contacts. This theoretical extreme we believe to be all but impossible. The mathematical model thus corroborates historical records documenting that in "virgin" populations, or in those of low or absent immunity, especially under crowded conditions, smallpox epidemics are deadly and brief. Information gathered at the time of the introduction of smallpox into Iceland in 1707, for example, supports this view; within the span of a few months smallpox killed about 18,000 of the country's 50,000 population, which was spread out over 40,000 square miles.[29] Similarly, in Aztec Mexico in 1520 between 3.5 and 15 million of 25–30 million persons, spread out over vast areas, are estimated to have died in an epidemic of less than six months.[30] Other examples of the devastation of smallpox in virgin and in relatively circumscribed populations abound: virtually all of these epidemics came and went in a matter of months. The possibility of a long duration smallpox epidemic in Athens is further confounded by the Potidaean expeditionary

---

[29]K. B. Roberts, "Smallpox, An Historic Disease," *Occasional Papers in Medical History* No. 1 (St. Johns, Newfoundland 1978).

[30]See J. A. Magner, *Men of Mexico* (Milwaukee 1942); F. F. Cartwright, *Disease in History* (London 1972).

epidemic, in which 1,050 of 4,000 hoplites died in about six weeks. Using the Maia model to describe the Potidaean outbreak would suggest an adequate contact number of 30–40 persons by "backwards" calculation. If exposure risks had been identical in the first Athenian epidemic (in fact, they may have been even greater, since unlike the Potidaean forces none could have been immune), a figure of only 20 adequate contacts would predict extinction of smallpox in less than eleven weeks. Because, however, there are no objective data allowing fair comparison of exposures during the original Athenian epidemic and the ensuing Potidaean outbreak, it is difficult to predict whether, and in what specific ways, crowding might have been worse in Athens than in military camps in the Potidaean countryside.

A crucial paradox to be confronted in all considerations of respiratory transmission is how a respiratory disease, explosive enough to devastate a hoplite expedition in six weeks, and a good part of the Athenian population in little more than that, can also be sufficiently indolent to linger within Athens/Piraeus for over two years, and perhaps for as long as five years. In an attempt to escape this paradox, we must also consider the possibility that the Athenian epidemic was caused by an explosive respiratory disease also capable of persistence via either traditional "reservoir" mechanisms, or by indolent rural transmission. Smallpox fits both of these criteria. Unlike most other respiratory diseases, viruses in dried smallpox secretions can survive for at least several months in clothes or bed linen, or even in such inanimate sources (fomites) as cotton bales,[31] suggesting a theoretical means of long term persistence of the agent beyond the chain of human–to–human transmission in Athens. So well documented is persistence in fomites that during the French and Indian wars Lord Jeffrey Amherst, Commander–in–Chief of British forces in North America, and Colonel Henry Bouquet apparently contrived to defeat Chief Pontiac's forces by giving them smallpox virus in contaminated blankets.[32]

It is also conceivable that after the lifting of the siege, smallpox was carried back to the countryside and then dispersed into a chain of indolent person–to–person transmission lasting as long as a year. Indolent smallpox transmission in nomadic or dispersed groups has been previously documented. Among North American Indian populations of the Great Lakes/St. Lawrence

---

[31]M. Young, "Small Pox Conveyed by Raw Cotton," *Br. Med. J.* 1 (1902) 687.

[32]J. Duffy, "Smallpox and the Indians in the American Colonies," *Bull. Hist. Med.* 25 (1951) 324–41. An early example of planned biological warfare, Bouquet wrote "I will try to inoculate the _____ with some blankets that may fall in their hands, and take care not to get the disease myself."

River area, introduction of smallpox in the early 1630s apparently led to regional transmission that continued for seven years.[33] Such indolent transmission also characterized the final days of smallpox during the worldwide eradication campaign in the 1970s. It is noteworthy that most such populations were at least partly immune, as returning populations of Atticans would surely have been. Shipboard smallpox transmission and survival over long ocean voyages is also well documented. As noted elsewhere, Thucydides implied that some other unspecified cities may have had the disease as well. It is thus conceivable that a combination of migrating Atticans and close allies, Athenians not initially infected, and susceptible Athenian expeditionary forces could constitute a critical mass of susceptible sufficient to sustain an epidemic surge following re-introduction from indolent foci. We believe this possibility warrants further scrutiny, including study of documented instances of indolent transmission of smallpox in peri-urban areas, and fitting of the Athenian data to more sophisticated mathematical models.

Existing data thus suggest to us that although purely respiratory diseases can be ruled out as causes of the Athenian epidemic, respiratory diseases capable of persistence in either focal reservoirs, or in fomites, or simultaneously explosive in crowded populations and capable of indolent transmission in dispersed rural populations, should remain under evaluation. Although most consistent with historically-documented smallpox, other long incubation diseases such as streptococcal diseases might also be candidates for further study with respect to mechanisms of persistence in rural and dispersed populations.

Reservoir transmission. Epidemiologic aspects of the Athenian epidemic appear to be most consistent with a disease associated with an infectious reservoir (an insect or animal vector). Of these, certain zoonotic diseases proposed as causes of the Athenian epidemic can be ruled out because human infection is accidental ("dead end") in the course of zoonotic transmission, and thus irrelevant to epizooticity. Among these are glanders, leptospirosis, rabies, and tularemia.[34] Psittacosis, which has not to our knowledge been previously suggested as a cause of the epidemic, can also be excluded on this basis.

---

[33]D. R. Hopkins, "A Destroying Angel," *Princes and Peasants. Smallpox in History* (Chicago 1983) 234–94.

[34]Glanders: C. H. Eby and H. D. Evjen, "The Plague at Athens: A New Oar in Muddied Water," *J. Hist. Med. Allied Sci.* 17 (1962) 258–63. Tularemia: J. A. H. Wylie and H. W. Stubbs, "The Plague of Athens: 430–428 B.C. Epidemic and Epizootic," *CQ* 33 (1983) 6–11.

Considered here under reservoir diseases, rather than under diseases of inoculation, anthrax[35] is an interesting possibility, primarily because by a combination of inhalation, inoculation, and ingestion, anthrax could be considered highly consistent with the clinical picture described by Thucydides. Not only are dogs affected by anthrax, but the organisms may be carried by birds, which are capable of contaminating water supplies. On clinical and zoonotic grounds anthrax is difficult to refute. But large–scale anthrax epidemics are not known to occur. What set of extraordinary circumstances might lead to epidemic anthrax is difficult to imagine. Secondly, there is no easily identified reservoir. Athenian sheep and cattle had been sent to Euboea, probably leaving only dogs, rats, mice, and birds within the city. An epidemic so massive and yet so chronic would have required a highly ubiquitous infectious source not only continually present, but exportable to the expeditionary forces. Spore–containing animal hides probably do not meet these criteria. Anthrax is thus difficult to reconcile with epidemiologic observations.

Other reservoir diseases suggested to have caused the Athenian epidemic include various insect–borne diseases, whether or not associated with zoonotic reservoir hosts, including malaria, plague, typhus, and various arboviral diseases, for example dengue, yellow fever, and Rift Valley fever.[36] All have been associated with explosive epidemics, and all except Rift Valley fever are closely linked to conditions of war, refugees, and overcrowding. Descriptive epidemiologic aspects of certain of the arboviral diseases, including dengue and Rift Valley fever, are probably more consistent with the Athenian epidemic than are epidemiologic aspects of any other disease. These arboviral diseases may not only be persistent, but re–emergent and explosive, satisfying the major observational criteria of Thucydides. The explosive behavior of Rift Valley fever has previously been noted. The explosive behavior of dengue has been repeatedly documented in both crowded and uncrowded situations, in times of

---

[35]F. Jahn, "Beiträge zur Geschichte Carbunkel–Krankheiten mit Ausschluss der Pest," *Janus Zeitschrift für Geschichte und Literatur der Medicin* (Breslau 1846) 369–414.

[36]Malaria: W. H. S. Jones, *Malaria and Greek History* (Manchester 1909) 33–40; Plague: F. Ranchin, "L'histoire de la Peste," *Opuscules ou traictez médecine divers et curieux en médecine* (Lyons 1640); A. Hirsch, "Beulenpest," *Handbuch der historisch–geographischen Pathologie. erste Abteilung. Infektionskrankheiten* (Erlangen 1859) 192–214; Typhus: Scarborough (above, n. 1) Arboviral diseases: in discussing arboviral diseases we refer to them generically rather than individually; comparisons to extant diseases would be misleading if, as appears likely, significant evolutionary changes in these agents had occurred over more than two millennia; Dengue: J.–P. Beteau, *La Peste d'Athènes (430–428 av. J.–C.) preface par M. le Docteur L. Tanon* (Paris 1934). Yellow fever: Smith (n. 17 above); Rift Valley fever: D. M. Morens and M. C. Chu, "The Plague of Athens," *New Engl. J. Med.* 314 (1986) 855.

war and of peace. For example, numerous explosive epidemics were recorded in the Pacific during World War II. Eighty per cent of U.S. Army troops stationed in Northern Territory and Queensland, Australia, got dengue in March–May 1942.[37] Because of troop rotations, the incidence rate in Espiritú Santo was 1,713 cases per 1,000 average troop strength in the month of April. The corresponding annual figure for Saipan in 1944 was 3,560 per 1,000. In peacetime, a 1977 Puerto Rican dengue epidemic swept through three million persons, with substantial population dispersion (only one thousandth as crowded as Athens in 430 B. C.), in five months.[38] Study of a 200–worker cohort during that epidemic revealed higher secondary attack rates in families (secondary attack rates are traditional indicators of respiratory spread) for mosquito–transmitted dengue than for respiratory–transmitted influenza, which was also epidemic at the time. Even in an open population of Puerto Rico (Bayamon municipio), dengue attack rates were paradoxically more highly correlated with family size than were influenza attack rates.

Also associated with dengue and other insect–borne diseases are door–to-door progression within neighborhoods, and increased risks for physicians and other persons caring for the ill. Had Thucydides observed the 1977 Puerto Rican epidemic of dengue and influenza cited above, he would probably have noted dengue's marked association with crowding, and perhaps even with infection of health care workers, before he noted the same for influenza, which is among the most explosive diseases transmitted by the respiratory route. In addition to a strong association with crowding, many arboviral diseases are also linked to water storage, particularly storage in cisterns (widespread in Piraeus) and in urns (widespread in Athens and Piraeus), where certain disease–causing peri–domestic mosquitoes selectively lay their eggs. In modern day Bangkok, explosive dengue epidemics may run through five million persons, spread out over hundreds of square miles in a period as short as two months, only to linger indefinitely, producing cases year–round, and to re-emerge periodically. Crowded areas with substantial water storage are usually stricken first and most dramatically, affecting multiple family members per house. The reservoir mode of transmission is consistent with the two to five year duration of the Athenian epidemic, with its explosiveness and its

[37]O. R. McCoy, "Dengue. Epidemiologic Considerations," in J. B. Coates, E. C. Hoff, P. M. Hoff, Medical Department, United States Army (eds), *Preventative Medicine in World War II, Volume VII: Communicable Diseases. Arthropodborne Diseases Other Than Malaria,* (Washington, D.C. 1964) IV 29–40.

[38]D. M. Morens, J. G. Rigau–Pérez, R. H. López–Correa, et al., "Dengue in Puerto Rico, 1977: Public health response to characterize and control an epidemic of multiple serotypes," *Am. J. Trop. Med. Hyg.* 35 (1986) 197–211.

association with crowding and contact exposure; it is the only obvious means of explaining why neither the Spartans nor the many other people with whom the Athenians came into contact over the five year period acquired the disease that was devastating tens of thousands of Athenians.

A serious problem in identifying any of the nearly countless vectorborne viral diseases, however, is that they evolve so rapidly: infectious agents existing 2,500 years ago would likely have undergone significant evolutionary change in the intervening centuries, and thus might not be recognizable today. Because, however, such evolution is usually most marked at the level of the genome, and least dramatic at the level of epidemiologic "behavior" (with changes in clinical appearance intermediate in degree), inferences about an arboviral cause of the Athenian epidemic that reflect epidemiologic observations are probably less liable to inaccuracies resulting from secular changes in the agent. Other than the arboviral diseases, malaria, plague, and typhus, have also repeatedly been associated with sudden explosive epidemics, followed by persistence of the causative agents over sufficiently prolonged periods of time to seed recurrent endemic or epidemic disease. Such other zoonotic viral agents as hantaviruses, arenaviruses, and (presumably) filoviruses also present clinical/epidemiological similarities, but because there is less available information about most of them, they are not considered further in this discussion.

Malaria is the least consistent of the remaining vectorborne diseases because of seasonal incompatibility and problems with epidemic explosiveness, particularly in closed populations such as Hagnon's expedition. Furthermore, because malaria was known to and recognized by Hippocrates and contemporary Athenian physicians, it is unlikely that Thucydides would have confused it with an epidemic disease considered to be novel. Finally, the clinical picture described by Thucydides is less consistent with malaria than are most other proposed vectorborne diseases.

Though consistent with the epidemiology of the Athenian epidemic, plague appears to be less consistent with the clinical description. Thucydides did not describe buboes, and deaths from pneumonic plague should have occurred much more rapidly (one to three days) than the disease Thucydides described.

Typhus, the classic explosive epidemic disease associated with wars, is consistent with the description of Thucydides, including the common occurrence of (actual) gangrene of the extremities, and of blindness. It may obviously persist in crowded populations for prolonged periods, and is also "persistent" in infected survivors who may, after long periods, have

recrudescences during which they again become infectious.[39] The chief argument against typhus is the ostensible presence of bullae in the Athenian epidemic. However, as noted above, it is not certain that Thucydides actually described bullae. In any case, vesiculobullous lesions have been described occasionally in typhus (MacArthur has claimed they are not uncommon in some epidemics,[40] although most experts would probably consider them rare), and their occurrence in rickettsialpox (an agent related to the agent of epidemic typhus) suggests that an ancestral rickettsial agent could have caused the Athenian epidemic. Examination of data from many typhus epidemics associated with wars and refugee situations suggests that the Athenian epidemic was typical of typhus in many other respects.

In summary, our epidemiologic evaluation of the Athenian epidemic excludes all common source diseases and most respiratory diseases. By a process of exclusion, the cause of the Athenian epidemic can be limited to either a reservoir disease (zoonotic or vectorborne), or one of the few respiratory diseases also associated with an unusual means of persistence: either environmental/fomite persistence, or adaptation to indolent transmission among dispersed rural populations. We suggest that the diseases in the first category include typhus, arboviral diseases, and plague, and in the second category smallpox. Both measles and explosive streptococcal disease appear to be less likely candidates, but historical and modelling research may serve to further characterize their suitability.

Our systematic approach to identifying the Athenian epidemic emphasizes descriptive epidemiologic methods, use of mathematical models, and fittings of the diseases' epidemiologic behavior to epidemics of known diseases documented in pre–modern times. We have de–emphasized reliance on clinical symptoms in favor of the disease epidemiology because pre–modern descriptions, which lack detailed information on serology and accurate accounts of rashes and other clinical features, will always retain a high degree of uncertainty. Although the framework presented here is both conceptual and preliminary, we suggest that there is much to be learned by comparing the Athenian epidemic to well–documented premodern epidemics of candidate diseases, such as those we cite. We believe that our approach has already limited the possible causes of the Athenian epidemic to a plausible few. Of these, typhus and smallpox may share the most clinical similarities with the

---

[39]For example, C. R. Green, D. Fishbein, I. Gleiberman, "Brill–Zinsser: Still With Us," *JAMA* 264 (1990) 1811–1812 report recrudescent typhus in a World War II concentration camp survivor more than 40 years after the infection.

[40]W. P. MacArthur, "The Athenian Plague: A Medical Note," *CQ* 48 (1954) 171–174.

Athenian epidemic. Since there exists an enormous body of empirically derived data on these candidate epidemic diseases in premodern times, we suggest as an avenue for future studies attempts to develop refined and situation–specific mathematical models for them using available historical information on their behavior in defined populations of varying sizes, crowding indices, immunities, and other relevant characteristics. In this manner, the limited number of remaining candidate diseases can be compared to the epidemic situation in Athens, in an attempt to find a "best fit". Such efforts will require an interdisciplinary approach that draws upon the talents of physicians, epidemiologists, biostatisticians, classicists, historians, and other scholars.[41]
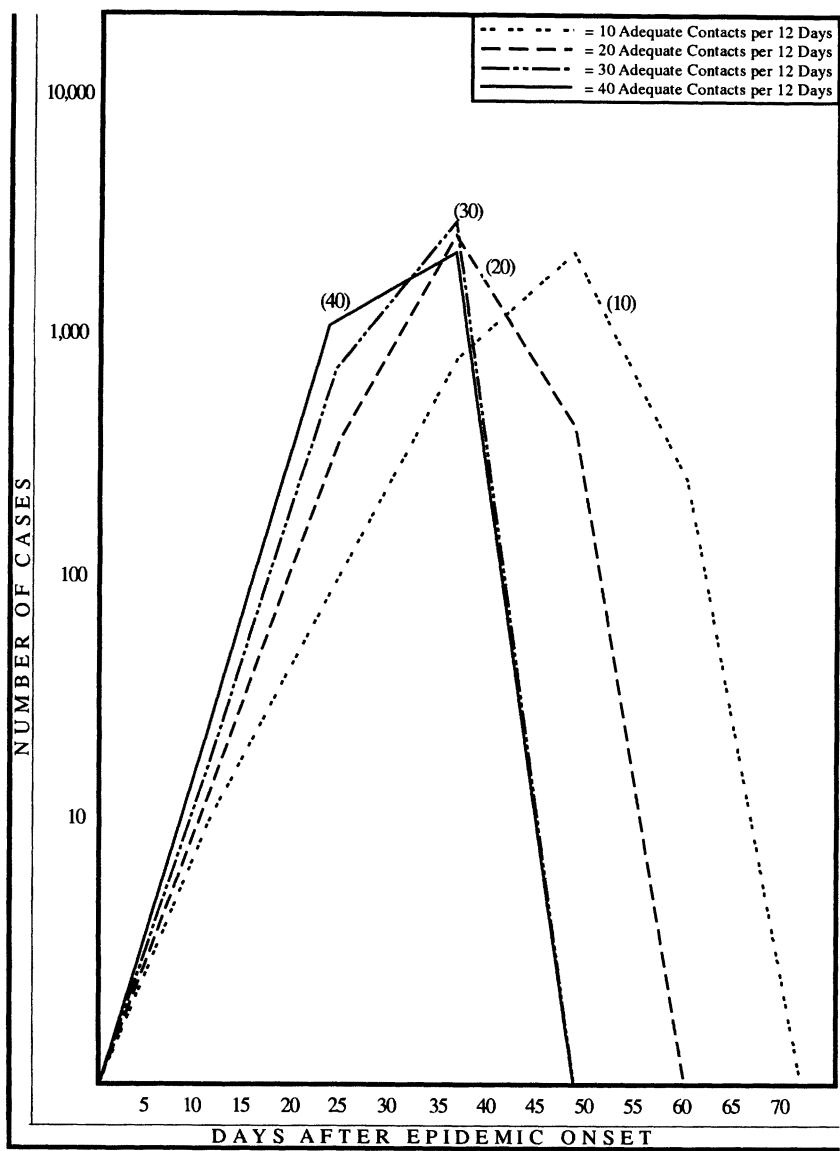
Figure 1. Theoretical epidemic curve of a transmissible respiratory disease of 12 day serial generation time (e.g. smallpox) in Hagnon's naval expedition to Potidaea in 430 B. C., a closed population of 4,000 persons, according to the deterministic mathematical model of Maia. Four separate curves are computed using adequate contact numbers of 10, 20, 30 and 40 per 12 days.
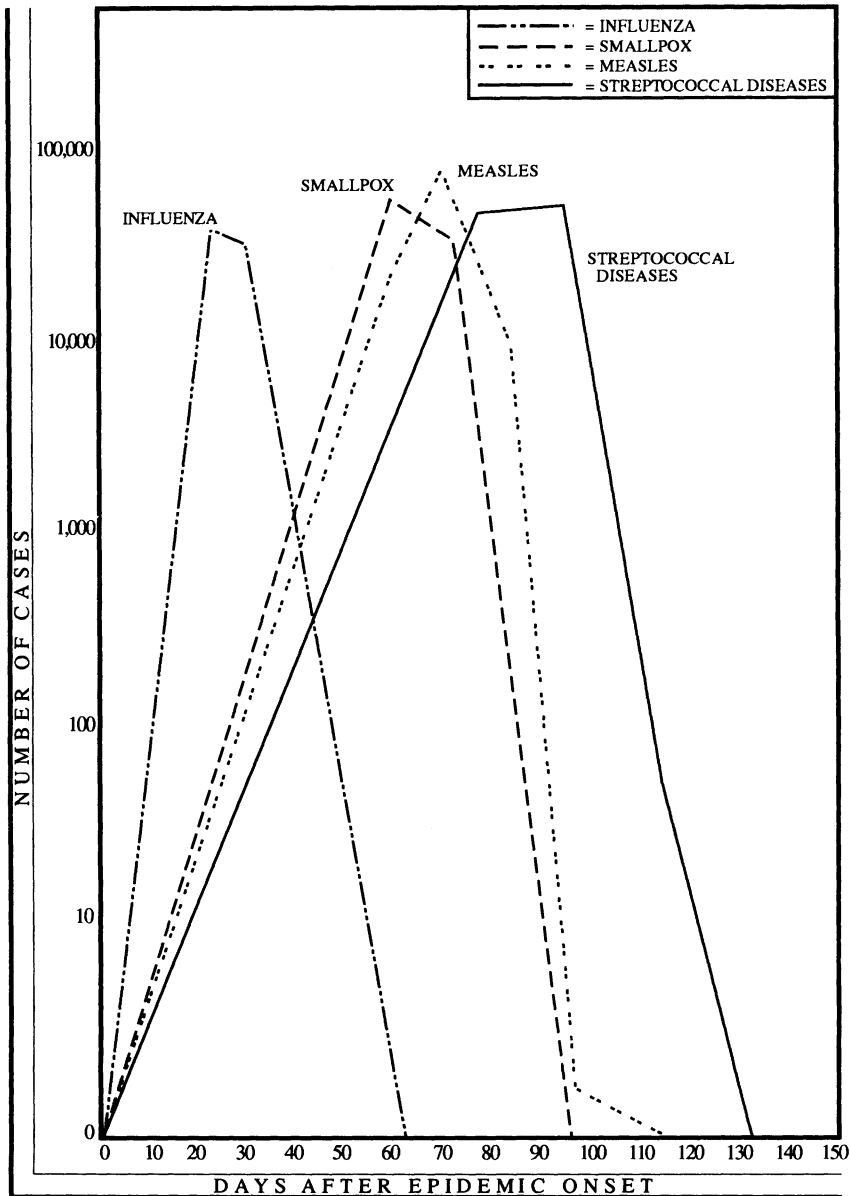
Figure 2. Theoretical epidemic curves of influenza A, smallpox, measles, and streptococcal disease in a closed population of 100,000 susceptible persons, according to the mathematical model of Maia, assuming an adequate contact number of 10.
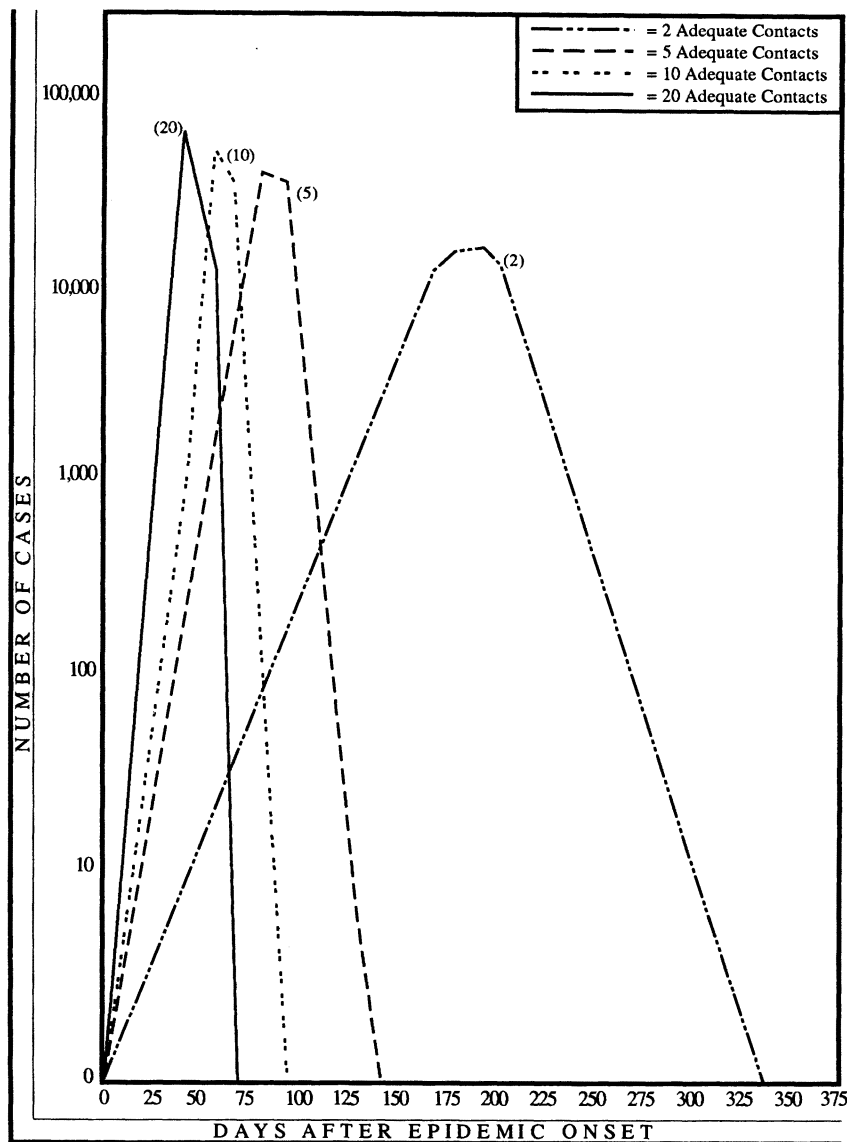
Figure 3. Theoretical epidemic curves of smallpox in a closed population of 100,000 susceptible persons, according to the mathematical model of Maia. The curves assume adequate contact numbers of 2, 5, 10, and 20 persons per serial generation time of 12 days.