International Journal of Astrobiology

cambridge.org/ija

Research Article

[†]The views expressed are those of the author and do not necessarily reflect the official policy or position of the California Military Department, the Air Force, the Department of Defense, or the US Government.

Cite this article: Sotos JG (2019). Biotechnology and the lifetime of technical civilizations. *International Journal of Astrobiology* **18**, 445–454. https://doi.org/ 10.1017/S1473550418000447

Received: 15 May 2018 Revised: 15 October 2018 Accepted: 5 November 2018 First published online: 15 January 2019

Keywords:

Drake equation; Fermi paradox; biotechnology; civilization lifespan; mathematical model

Author for correspondence: John G. Sotos, E-mail: ija2019@sotos.com

© Cambridge University Press 2019. This is a work of the U.S. Government and is not subject to copyright protection in the United States.

CAMBRIDGE UNIVERSITY PRESS

Biotechnology and the lifetime of technical civilizations

John G. Sotos[†]

Air Division, Joint Forces Headquarters, California National Guard, Sacramento, CA 95826, USA

Abstract

The number of people able to end Earth's technical civilization has heretofore been small. Emerging dual-use technologies, such as biotechnology, may give similar power to thousands or millions of individuals. To quantitatively investigate the ramifications of such a marked shift on the survival of both terrestrial and extraterrestrial technical civilizations, this paper presents a two-parameter model for civilizational lifespans, i.e. the quantity L in Drake's equation for the number of communicating extraterrestrial civilizations. One parameter characterizes the population lethality of a civilization's biotechnology and the other characterizes the civilization's psychosociology. L is demonstrated to be less than the inverse of the product of these two parameters. Using empiric data from PubMed to inform the biotechnology parameter, the model predicts human civilization's median survival time as decades to centuries, even with optimistic psychosociological parameter values, thereby positioning biotechnology as a proximate threat to human civilization. For an ensemble of civilizations having some median calculated survival time, the model predicts that, after 80 times that duration, only one in 10²⁴ civilizations will survive - a tempo and degree of winnowing compatible with Hanson's 'Great Filter.' Thus, assuming that civilizations universally develop advanced biotechnology, before they become vigorous interstellar colonizers, the model provides a resolution to the Fermi paradox.

Introduction

In 1961 Drake introduced a multi-parameter equation to estimate the number of civilizations in the galaxy capable of interstellar communication (Drake 1961). Soon after, von Hoerner (1961) and Shklovskii and Sagan (1966) concluded that the equation's precision depended principally on its parameter L – the mean lifetime of a communicating civilization – because L's value was uncertain over several orders of magnitude. While subsequent advances in astrophysics have improved the precision of several parameters in the Drake equation (Burchell 2006; Frank and Sullivan 2016; Vakoch and Dowd 2015), L remains highly uncertain (Oliver and Billingham 1971; Ambartsumian and Sagan 1973; Billingham *et al.* 1979; Duncan 1991; Schenkel 1999; Kompanichenko 2000; Rubin 2001; Forgan 2009; Maccone 2010).

The apparent absence of communicating civilizations (Webb 2015) in our planet-rich galaxy (Cassan *et al.* 2012) underscores the possibility that such civilizations have short L (Webb 2015; Bostrom and Cirkovic 2011), potentially due to factors exogenous to the civilization (e.g. nearby supernovae) and/or endogenous to the civilization (e.g. self-destruction).

On Earth, control of endogenous factors that could destroy civilization – namely, Malthusian resource exhaustion, nuclear weapons and environmental corruption – has until now rested with the very few persons who command large nuclear arsenals or steer the largest national economies. However, emerging technologies could change this. For example, biotechnology (President's Council of Advisors on Science and Technology 2016) and nanotechnology (Drexler 1987) offer the prospect of self-replicating elements able to spread autonomously and calamitously worldwide, at low cost and without heavy industrial machinery. Ultimately, thousands of individuals – having varying levels of impulse control – could wield such technologies.

Intuition suggests danger rises as potentially civilization-ending technology ('CE technology') becomes more widely distributed, but quantitative analyses of this effect in the context of Drake's L are rare. At the extreme of technology diffusion, Cooper (2013) modelled an entire population of 10^{10} individuals (growing at 2% annually), each with a 10^{-7} annual probability of unleashing a biological agent causing 50% mortality (with 25% standard deviation). He found a mean span of L=8000 years before extinction, defined as a population less than 4000.

¹For brevity, 'civilization' in this paper refers to a civilization capable of interstellar communication, and the 'lifespan' or 'lifetime' of a civilization is the span of time during which it is able to communicate. Thus, the 'death,' 'silencing,' or 'ending' of a civilization are synonymous.

This paper generalizes Cooper's work. It develops a simple two-parameter mathematical model for L that applies to most scenarios of disseminated CE technology and is mathematically indifferent to specific CE technologies. For reasons summarized below, however, biotechnology may be regarded as a universal CE technology.

Biotechnology's potential to end civilizations

On Earth, microbial pandemics have ended non-technical civilizations (McNeill 1976). Antimicrobial drugs mitigate such risks only partially. Advisors to the President of the USA have already warned that biotechnology's rapid progress may soon make possible engineered microorganisms that hold 'serious potential for destructive use by both states and technically-competent individuals with access to modern laboratory facilities' (President's Council of Advisors on Science and Technology 2016). Indeed, small research groups engineered proof-of-principle demonstrations years ago (Jackson et al. 2001; Herfst et al. 2012; Imai et al. 2012), while recent history provides a precedent not only for a laboratory-preserved organism causing a worldwide pandemic² (Wertheim 2010; Rozo and Gronvall 2015), but also for the organism's descendants circulating for 30 years in the global population (Zimmer and Burke 2009). Looking forward, medical research initiatives such as the Cancer Moonshot (National Cancer Institute 2018) may, if successful, seed thousands of hospitals with exquisitely targetable cell-killing biotechnology that could, in principle, be adapted and aimed at any genetically defined target, not just cancer cells.

Any technically-capable intelligence produced by evolution likely shares this susceptibility. 'Genetic' processes, defined here as those that pass information to build a succeeding generation or direct the self's use of sustaining energy, are required for evolution (Farnsworth et al. 2013). Assuming that no process can be perfect, imperfections in genetic processes equate to 'genetic diseases,' and will spur any intelligence having selfpreservation drives to develop genetic manipulation technology to ameliorate those diseases. Given this motivation to alter genetic processes, plus the biological certainty that genetic processes respond to environmental inputs (e.g. food shortages), plus a general technical capacity to control environments ever more precisely, the eventual appearance of biotechnology may be expected. Cooper (2013) expects that civilizations will typically develop biotechnology and spaceflight approximately simultaneously.

Biotechnology is inescapably threatening because it is inherently dual-use (Watson *et al.* 2018): curing genetic disease enables causing genetic disease. Cooper (2013) uses Cohen's (1987) theorem to assert that, under any reasonable model of computing (applied here to bio-molecular computing), no algorithm ('medical treatment') can stop every possible piece of invasive self-replicating software. Whether Cohen's theorem strictly applies or not, the truism that defensive technology generally lags offensive is relevant.

Of course, any civilization can walk away from any technology. But, because other widely available technologies with civilization-ending potential, e.g. nanotechnology, lack the *a priori* universal desirability of biotechnology, only biotechnology will herein be further discussed.

Model and results

The baseline model assumes that all communicating technical civilizations either continue communicating forever or go silent involuntarily due to some action arising within each civilization. Two parameters model the lifespan of such civilizations: E, the number of entities (individuals, coalitions, nation-states, etc.) in the civilization who control a means to end civilization (i.e. render it uncommunicative), and P, the uniform probability per annum per entity that an entity will trigger its civilization-ending means. Entities act independently, and civilization is assumed to end with the first trigger.

The simplest model for the probability, C(y), that the civilization will still be communicative after y years, under constant E and P, is:

$$C(y) = (1 - P)^{(Ey)} \tag{1}$$

Solving equation (1) for y:

$$y = \frac{\ln C(y)}{E \ln(1 - P)} \tag{2}$$

Borrowing the abbreviation LD_{50} from pharmacology, where it indicates the median lethal dose of a substance, it is here re-conceptualized as 'lethal duration 50' to indicate the number of years, under a given E and P, before civilization's accumulated probability of being uncommunicative, 1 - C(y), is 50%. Substituting C(y) = 1 - 0.50 into equation (2) yields:

$$LD_{50} = \frac{\ln(1 - 0.50)}{E \ln(1 - P)} \tag{3}$$

Similarly, the number of years before civilization has a 5% chance of becoming uncommunicative is:

$$LD_{05} = \frac{ln(1 - 0.05)}{E ln(1 - P)} \approx (0.074 \ LD_{50}) \approx \frac{LD_{50}}{13.5}$$

Increasing the certainty of civilizational death increases the lethal duration exponentially, as Fig. 1 shows. Thus, for any E and P, $LD_{95} \approx (4.3 \ LD_{50})$, $LD_{99,9999} \approx (20 \ LD_{50})$, and $LD_{100[1-C(y)]} \approx (80 \ LD_{50})$ where $C(y) = 10^{-24}$.

Figure 2 plots the relationship between E and LD_{50} for several P, and illustrates the approximation $LD_{50} \approx 0.7/(E \times P)$, derived in equation (A3) of the Mathematical Appendix.

To calculate the mean lifespan, it is more intuitive to first calculate the number of communicating civilizations, N(w), that exist at the end of a time window extending from year y = 0 to y = w. Assuming that zero civilizations existed at y = 0, and that communicating civilizations were born at a constant rate of B per year throughout the time window, the Mathematical Appendix shows:

$$N(w) = B \int_0^w C(y) \, dy$$
 (is A5)

$$= B \frac{S^{w} - 1}{\ln S}$$
 where $S = (1 - P)^{E}$ (is A8)

²This pandemic miserably sickened the author in early 1978.

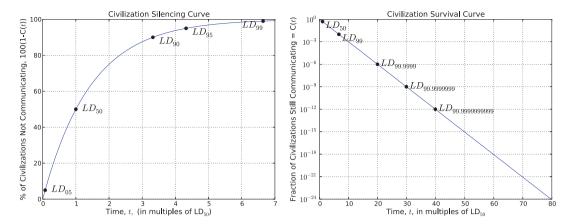


Fig. 1. Survival times in a cohort of civilizations, all created at t=0. Left: Over time, the percentage of silent civilizations, 100(1-C(t)), logarithmically approaches 100%. For any E and P, $LD_{X\%} = (ln(1-X\#))/(ln(1-0.50))LD_{50}$, where X% is a percentage and X# is the equivalent probability. Right: This panel modifies the left panel's axes. First, the time axis is expanded compared with the left. Second, the vertical axis has been inverted to show survival, C(t), over time. The LD_{50} and LD_{99} points carry over from the left panel. Remarkably, the time required to reach infinitesimal survival rates, e.g. 10^{-24} , is less than two orders of magnitude larger than the median civilizational survival time, LD_{50} .

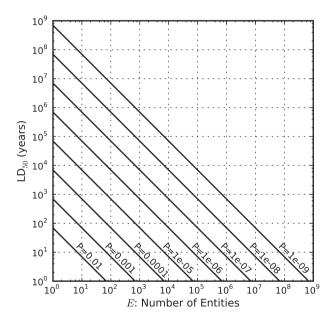


Fig. 2. Technology diffusion (E) and psychosociology (P) determine the civilizational lifespan (LD_{50}). E is the number of entities who control a means to end civilization. P is the probability per annum per entity that the entity will trigger its civilization-ending means. Given a constant E and P, LD_{50} is the median number of years before civilization is expected to end. E and LD_{50} have an inverse linear relationship for any P.

$$\approx \frac{B}{EP} \left[\text{for } w \ge \frac{10}{EP} \text{ and small } P \right]$$
 (is A10)

Figure 3 plots the exact form of N(w) from equation (A8), for multiple w and EP when B=1. It shows with reasonable precision that $N(w) \le B/(EP)$ for any w.

The parameter L in the Drake equation is reformulated herein to $\overline{L(w)}$, the mean lifespan for civilizations born during a time window of duration w. This transforms the Drake equation to:

$$N(w) = B\overline{L(w)}$$
 (is A12)

Thus, $\overline{L(w)} = N(w)$ when B=1, and so Fig. 3 is also a plot of $\overline{L(w)}$.

Per Fig. 3, $\overline{L(w)}$ increases with w. However, its maximum value, at any time, is constrained. Assuming all civilizations have identical E and identical E:

$$\overline{L(w)}_{\text{max}} < \frac{1}{EP}[\text{for all } w]$$
 (is A15)

Combining these two formulae and defining N as 'N(w) for all w' yields the Drake equation as an inequality:

$$N < \frac{B}{FD} \tag{4}$$

or, hewing to its classical form (Drake 1961):

$$N < \frac{R_* f_p n_e f_l f_l f_c}{FP} \tag{5}$$

Because the model addresses only endogenous involuntary silencings, adding consideration of other causes for silencings would merely reinforce this inequality.

To produce near-term risk estimates for Earth, a PubMed search informed the value of E, as follows. With the assumption of a civilization-ending technology based on some yet-to-be-described genetic technique, the number of people authoring scientific articles indexed under 'genetic techniques' (one of PubMed's \approx 27 000 standard index terms) can be used to estimate the number of people capable of exploiting such a technique, thereby serving as a proxy for E. Thus, the PubMed search

performed on 10 August 2017, yielded 594,458 publications in the most recent 8-year span of complete bibliographic coverage. After eliminating non-scientific publications (of type letter, comment, news, interview, etc.) 585 004 remained, which carried 1 555 661 unique author names. Of these authors, approximately 179 765 appeared on five or more publications. This number is a maximum because some authors publish under more than one name.

Models employing non-constant E and P are possible. The simplest posits that E grows as a population might: a fixed percent

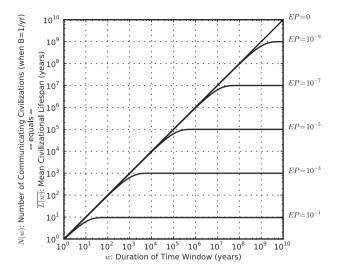


Fig. 3. Civilizations and time. For six different values of $E \times P$, the plot shows two equivalent quantities for time windows of various durations w: (a) $\overline{L(w)}$ = mean lifetime of communicating civilizations over time, and (b) N(w) = number of communicating civilizations over time when B=1. For both quantities, a constant B is assumed. Zero civilizations exist at time w=0. Equation (A15) mandates $\overline{L(w)} < 1/(EP)$ for all w. Per equation (A9), $\overline{L(w)}$ grows substantially until a near-steady state is reached at about w=10/(EP) years. An arbitrary-precision software package (Johansson et al. 2013) used equation (A8) to calculate N(w) and $\overline{L(w)}$.

per year. If, over y years, E grows this way from some initial value E_0 , with the growth continuously compounded, then:

$$E_{y} = E_0 e^{ry} \tag{6}$$

where r is the growth factor (e.g. 0.02 for 2% annual growth) and e = 2.71828... Unfortunately, the unbounded exponential term renders this 'growth model' nonsensical for even moderately large y. Still, some insights can emerge for short-time horizons, as detailed in Fig. 4, which is based on equation (A18) in the Appendix. Unsurprisingly, a growing E yields an LD_{50} significantly smaller than is calculated from a constant E.

Discussion of model

Unless explicitly noted, all discussion refers to the baseline model in which *E* is constant.

Equation 1 provides the probability, C(y), that a civilization survives endogenous involuntary silencing threats until some L = y. The lethal durations, LD_{50} et al, are probabilistic statements of this L. Because the model terminates upon the first use of a civilization-ending technology, more complicated models, such as the Poisson distribution, are not required.

Thus, the model is simple, but not unreasonably so. However, with only two parameters, it is important to understand their inherent assumptions.

Model discussion: E and P (and B)

In broad terms, *E* characterizes a CE technology and its availability, while *P* characterizes the psychology and sociology of the entities who possess the technology. Although the loss of interstellar communicativeness is equated to the end of civilization, other endpoints (e.g. complete extinction) could be substituted.

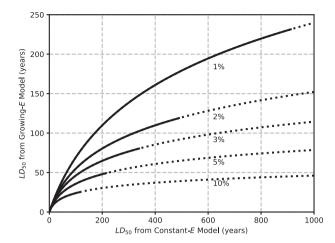


Fig. 4. Drop in LD_{50} when E grows. The horizontal axis corresponds to LD_{50} values calculated from equation (3) and a constant E and P. If, however, E is not constant, and instead grows at a fixed percentage annually (five growth rates are shown), then LD_{50} shrinks to the corresponding value on the vertical axis, according to equation (A19). So, for example, an LD_{50} of 600 years derived from Fig. 2 would be revised to approximately 190 years if E grew by 1% annually. To signal wariness about exponential explosion, each solid line changes to a dotted line when the number of entities has increased a million-fold (i.e., $E_V/E_0 \ge 10^6$).

The only criteria are consistency of the endpoint, independence of the entities, and termination of the model upon first triggering. Numerous subtleties attend the definitions of E and P.

First, *E* applies to any CE technology, be it nuclear, nano-, bio-, or another. The CE technology never fails to end civilization once triggered. The effects of 'near miss' extinction events on population and psychology are ignored.

Second, E includes only entities that possess (or can acquire) the 'full stack' of CE technology. That is, they must have the capability to make or otherwise obtain the weapon, and to deliver it in quantities that render the civilization uncommunicative. So, for example, even though designs for nuclear weapons are comparatively well known (Phillips 1978), E for Earth remains only ≈ 2 (representing the leaders of the USA and Russia). The self-propagating nature of biological weapons would simplify, but not eliminate, the delivery challenge.

Third, to the extent that machine intelligences possess CE technology, they could also be counted in *E.* (Exemplar: 'SkyNet' from the *Terminator* movies.)

Fourth, *E* reflects a balance between offensive and defensive technologies. Thus, developing and readying defensive technology offers a straightforward, albeit challenging, path to markedly decrease *E*.

Fifth, *P* is the sum across all reasons, intended or not, that an entity might trigger the CE technology. Most are psychosocial, e.g. greed, hate, stupidity, folly, gullibility, power-lust, mental illness, ineptitude, non-fail-safe design, etc. *The Bulletin of the Atomic Scientists*' 'doomsday clock' (Anonymous 2002) has similarities to *P*.

Sixth, the model assumes constant *E* and *P* throughout the time window of interest. This is unlikely to occur in a real civilization, given the dynamics of offensive/defensive technologies, population, sociopolitical stability, and technology diffusion. Simple model extensions would have *E* and *P* vary over time, or

 $^{^3}E$ would be slightly higher if additional other leaders could end civilization via climate change.

sum across subpopulations of entities each with their own E_i and P_i , or sum across multiple CE technologies each with their own E_j and P_i .

Unlike the model of (Cooper 2013), population growth – and the concomitant growth in E – is omitted from the baseline model because all realistic non-zero growth rates become nonsensical when compounded (exponentiated) over eons. Over short time-frames, the effect of a growing E can be reasonably equated to a speed-up in time. For example, when E is constant and the model reaches some state at year y, a situation in which E is growing by 2% annually will attain the same state significantly earlier, at year $50 \ln (1 + y/50)$ according to equation (A19).

The model's flexibility could be improved – at the cost of great mathematical complexity – by assigning probability distributions to E and P and convolving them. However, models that assume a distribution around some mean value for P (denoted $P_{\rm mean}$) will yield lower values for C(y) and LD_{50} than the present model, because of the positive exponent in the definition of C(y). Thus, this model's dispiritingly low values for LD_{50} nevertheless represent a civilization's best-case outcome for a given $P_{\rm mean}$.

This is most obviously appreciated in the edge case where a single entity has its P=1, for example, an entity who acquires the skills of a CE technology specifically to end civilization. As soon as a single qualified entity has P=1, then the overall civilizational P is also 1, and LD_{50} (in fact, all LD_x) is zero.

Civilizations spanning multiple planets should be treated as multiple civilizations, each modelled separately with their own E and P. Modelling them as a single civilization assumes all the planets' civilizations die from one attack – an unnecessarily stringent requirement. Of course, P might change on planets that see a sister planet destroy itself.

Although colonization would imply a non-constant B, the model would still apply so long as B is less than some constant B_{max} . Using B_{max} in the model would provide an upper bound for N(w). Geometrically increasing B would require re-working the model, but the barrenness of the galaxy mitigates this possibility: (Tipler 1980) and others (Webb 2015; Jones 1981; Armstrong and Sandberg 2013) note that a single civilization colonizing at even moderate rates of geometric increase would fill the galaxy in only a few million years, and we do not observe a full galaxy.

Furthermore, assuming that the technology of interstellar colonization is far more daunting than biotechnology, and that the self-preservation drives of individual intelligences far exceed any elective desire to migrate off-planet, it is reasonable to expect that, as a rule, civilizations will develop and use sophisticated biotechnology before dispersing themselves on other planets (Cooper 2013). Thus, the experience of 20th century Earth is likely typical, i.e. the progress of medicine and public health in the era antedating genetic biotechnology creates a population explosion, so that civilization consists of a large, dense, mobile population on a single homeworld at the time that potentially CE biotechnology is developed. Because such ecological conditions are conducive to the spread of communicable agents, it is reasonable to hypothesize that all planetary civilizations will face existential threats from contagious micro-organisms - whether engineered or not - before they become vigorous interstellar colonizers (Cooper 2013).

⁴The model would become complex to the extent that interaction terms would be needed to model a single entity having access to multiple CE technologies. However, modeling the technologies separately and then choosing the most pessimistic outcome would likely suffice.

The model could also apply to civilizations based on networked machine intelligences when epidemic malware is a possibility. Because diversity among evolution-produced organisms would likely be higher than among designed software, building CE technology against machine intelligences could be comparatively easy.

Model discussion: stability

It may be argued that a potential CE technology cannot exist for long time spans without a defensive technology being developed, i.e. that *E* cannot exceed zero for thousands, millions, or billions of years.

Several considerations weaken this proposition, especially as relates to biotechnology. These considerations are illustrative and necessarily speculative. Future biotechnological progress will elucidate the extent to which they hold.

First, reliance on a single CE technology is not required. Instead, multiple CE technologies may exist serially, each enabling a multitude of different attacks, with each attack requiring a different defense. This is akin to the inventory of 'zero day exploits' that present-day entities accumulate to penetrate computer systems.

Second, a long period of E>0 can be viewed as the concatenation of shorter time periods having $E_i>0$, where each E_i derives from a separate CE attack possibility that is eventually countered by a defense tailored to that attack. For example, if the frailties of life allow for a million different attacks,⁵ and they are arrayed sequentially, and it takes 1 year to tailor a defensive technology for each, then E>0 for $w=10^6$ years. If no periods of E=0 were interspersed between the $E_i>0$ periods, then the time window w would equal elapsed time in the universe. In scenarios having interspersed $E_i=0$ periods, elapsed time would exceed window duration.

Third, the mere development of defensive technology is not sufficient. The technology must be fully fielded. That is, unless widespread pre-exposure vaccination is possible, an attack must be detected, the agent(s) characterized, and the remedy developed, tested, manufactured (perhaps in billions of doses), distributed and administered – all of which must succeed before the attack can take root in the population. This is a formidable challenge requiring multiple sub-technologies in the near term, or a single future technology that is currently indistinguishable from magic.

Fourth, defensive technology may be impossible on first principles. For example, every known life form adapts its gene expression to its environment. An offensive technology whose only defense necessitated extinguishing this genetic responsiveness would seem unobtainable.

Fifth, mere possession of defensive technology is not sufficient – timely and correct decisions to activate defenses on a civilizational scale must also occur. Thus, a civilization's decision-making process, be it political, machine-based, or other, is also a target for CE technologies. This means *E* has a small psychosociological component. Decentralized decision-making, such that

 $^{^5}$ Even simple viruses have profound combinatorial reserve. Influenza A, for example, with its genome of $\approx 14\,000$ nucleotides, has ≈ 880 million combinatorial two-nucleotide variants and ≈ 12 trillion three-nucleotide variants (Perelson *et al.* 2012). Though only a sliver of these would yield functionally and/or immunologically distinct viruses, the numerator explodes exponentially. It is a tall order to devise anti-influenza A technologies that are 100% effective against all possible variants.

⁶Alternatively, the model could divide P into P_{offense} and P_{defense} .

every individual intelligence possessed the counter-CE technology and independently decided when and if to self-medicate, would require a level of trust in the population that no government on earth has so far developed.

Sixth, generalizing the above scenario, CE technologies need not be highly lethal. To sustain itself, a densely populated world may rely on critical infrastructure and/or heavily optimized industrial processes. Direct or indirect disruption of these essential functions could cause sufficient social chaos to render a civilization uncommunicative.

Finally, if *EP* is large throughout the universe, then the model does not have to apply for millions or billions of years. For example, if $E = 10^3$ and $P = 10^{-3}$ then $LD_{50} \approx 0.7$ years and the probability of surviving to 25 years is $< 10^{-9}$.

Discussion of results

Results discussion: Earth

From equation (A3), achieving $LD_{50} \ge 1000$ years requires $EP \le 7 \times 10^{-4}$. Thus, with E = 2 today, $P \le 0.00035$ is required.

Given the pace of biotechnology's progress, plus the irresistible pressure to continue that progress for universally-desired medical purposes, plus the dual-use potential of the technology, plus its potential worldwide reach, many humans could soon have the capacity to end Earth's technical civilization, driving $E\gg 2$. In a recent 8-year span, more than 1.5 million people participated in the 'genetic techniques' enterprise at a level sufficient to warrant authorship on a scientific article. Almost 180 000 of them authored five or more such articles. The number actually engineering artificial organisms today is certainly far smaller, but clearly, a large reservoir of hands-on molecular genetics competence already exists on Earth.

Although LD_{50} has been our focus, planning with lower thresholds (Suskind 2006), e.g., LD_{05} ($\approx LD_{50}/13.5$) or LD_{01} ($\approx LD_{50}/70$), would mitigate unanticipated rapid rises in E or P. For example, comparing a CE technology's LD_{01} to the anticipated time needed to develop defensive counter-technology might drive policymakers to speed such development.

Given the PubMed authorship numbers, a few new biotechnological innovations could reasonably and quickly raise E to 10^4 . If so, and $P=10^{-7}$, then $LD_{01}\approx 10$ years. If E became larger, LD_{01} would become smaller. The short LD_{01} time span is concerning, given today's comparatively slow pace of antimicrobial innovation (the common cold and many other infections remain incurable and without vaccinations), and strongly argues that defensive technology development must be expanded and must occur simultaneously with any therapeutic (offensive) development.

An especially concerning scenario arises if, someday, hospitals employ people who routinely write patient-specific molecular-genetic programmes and package them into replicating viruses that are therapeutically administered to patients, especially cancer patients. If the world attained the European Union's per capita hospital density, this could mean two hundred thousand hospitals employing perhaps 1 million people who might genetically engineer viruses every workday. Should techniques emerge for a highly communicable therapeutic virus – against which

vaccination would be refused, as that would preclude future cancer therapy – and E reached 10^6 , then attaining an LD_{01} of just 10 years would require $P < 10^{-9}$, perhaps an impossibility, given human nature.

Results discussion: Drake equation

By simulating an ensemble of civilizations, the present model challenges Burchell's (2006) assertion that L in the Drake equation is 'not truly estimable [estimatable] without observation of a set of societies.' Although estimating P based on first principles cannot be done for extraterrestrial civilizations, estimating E and the product EP may be tractable within the assumptions of the model, as follows.

Lower-bound estimates for E would derive from deep understanding of the genetic mechanisms of life – all possible mechanisms, not just DNA/RNA – and from the possibilities of biotechnology as applied to those mechanisms. Thus, estimates of E would derive from understanding the gamut of intelligence-compatible biologies, an understanding that smart human biochemists could perhaps achieve ex nihilo, without interstellar travel or communication. Machine intelligences would have analogous considerations. The existence of other CE technologies might increase E further.

Because of equation (4), EP can be constrained by searching for extraterrestrial intelligence (SETI). With B increasingly well understood, constraining N in equation (4) constrains EP. Thus, if SETI efforts someday yielded a conclusion such as 'We estimate that no more than N_x communicating civilizations exist,' then $EP < B/N_x$.

If both EP and E can be estimated, then the value of P is constrained. It is interesting to note that, given its dependence on psychological factors, possessing a constraint or estimate of P would be the first step toward a quantitative epidemiology of alien psychologies.

The model applies so long as opportunities to deploy civilization-ending means predate the ability to counter all such attacks (and accidents). That is, whenever E > 0, equation 3 produces a finite value for LD_{50} and civilization is at risk, assuming P > 0. Whether any measures could achieve P = 0, short of pervasive and perfect surveillance of entities, is unknown.

The model's low values for lifespan, $\overline{L(w)}$, have implications for SETI strategy. If geometrically increasing interstellar colonization circumvents short civilizational lifespan, then, all other factors being equal, communicating civilizations would be longest-lived where such colonization is easiest, e.g. where the time and/or energy required to move between habitable planets is smallest. This consideration adds to existing reasons why SETI might target zones of densely collected habitable planets (Turnbull and Tarter 2003).

Results discussion: the Fermi paradox and the great filter

To date, in a visible universe of $\approx 10^{24}$ stars and their planets, only Earth shows evidence of intelligent life. This apparent paradox, noted by Enrico Fermi and others (Webb 2015), could be explained by a 'Great Filter' that all but prevents communicating civilizations from forming or surviving (Hanson n.d.). The Great Filter may be technological in origin if '(a) virtually all sufficiently advanced civilizations eventually discover it and (b) its discovery leads almost universally to existential disaster' (Bostrom 2008).

⁷In 2004, 15 000 hospitals (European Hospital and Healthcare Federation 2009) were serving 500 million people (Organisation for Economic Cooperation and Development European Union 2016). Likely, few emerging health systems will follow an American model.

Most remarkably, the present model supplies the quantitative 24 orders-of-magnitude winnowing required of a Great Filter, reducing it to a two-orders-of-magnitude multiplication. For example, if $E=10^6$ and (optimistically) $P=10^{-9}$, then $LD_{50}\approx 700$ years, and $LD_{100[1-C(y)]}\approx (80\ LD_{50})\approx 56\,000$ years when $C(y)=10^{-24}$. That is, for this E and E, we expect only one civilization in E 100 000-year lifespan is effectively impossible because only one in E 100 000-year lifespan is effectively impossible because only one in E 100 000-year lifespan is communicative.

Overall, therefore, I would advise advanced technical civilizations to optimize not on megascale computation (Sandberg *et al.* 2017) nor engineering (Dyson 1960) nor energetics (Kardashev 1964), but on defense from individually-possessable self-replicating existential threats, such as microbes or nanomachines.

Author ORCIDs. (D) John G. Sotos 0000-0003-0176-2907.

Acknowledgments. I am grateful for the support and wise counsel of Jennifer Esposito, Mike Morton, Barry Hayes, and, of course, Tanya Roth. However, all errors are the author's responsibility. My thanks go to the anonymous reviewer for spurring development of Fig. 4.

References

- Ambartsumian VA and Sagan C (1973) Prospect. In Sagan C (ed.), Communication with Extraterrestrial Intelligence. Cambridge, MA: MIT Press, pp. 1–7.
- **Anonymous** (2002) The history of the Bulletin clock. *Bulletin of the Atomic Scientists* **58**(2), 36–37. (http://dx.doi.org/10.1080/00963402.2002. 11460553)
- Armstrong S and Sandberg A (2013) Eternity in six hours: intergalactic spreading of intelligent life and sharpening the Fermi paradox. Acta Astronautica 89, 1–13. (http://www.sciencedirect.com/science/article/pii/S0094576513001148)
- **Billingham J, Oliver BM and Wolfe JH** (1979) A review of the theory of interstellar communication. *Acta Astronautica* **6**(1), 47–57.
- **Bostrom N** (2008) Where are they? Why I hope the search for extraterrestrial life finds nothing. *Technology Review* 72–77. (https://www.technologyreview.com/s/409936/where-are-they/)
- **Bostrom N and Cirkovic MM** (2011) *Global Catastrophic Risks*. Oxford: Oxford University Press.
- Burchell MJ (2006) W(h)ither the Drake equation?. International Journal of Astrobiology 5(3), 243–250.
- Cassan A, Kubas D, Beaulieu J-P, Dominik M, Horne K, Greenhill J, Wambsganss J, Menzies J, Williams A, Jorgensen UG, Udalski A, Bennett DP, Albrow MD, Batista V, Brillant S, Caldwell JAR, Cole A, Coutures C, Cook KH, Dieters S, Prester DD, Donatowicz J, Fouque P, Hill K, Kains N, Kane S, Marquette J-B, Martin R, Pollard KR, Sahu KC, Vinter C, Warren D, Watson B, Zub M, Sumi T, Szymanski MK, Kubiak M, Poleski R, Soszynski I, Ulaczyk K, Pietrzynski G and Wyrzykowski L (2012) One or more bound planets per milky way star from microlensing observations. Nature 481(7380), 167–169. (http://dx.doi.org/10.1038/nature10684)
- Cohen F (1987) Computer viruses: Theory and experiments. Computers & Security 6, 22–35.
- Cooper J (2013) Bioterrorism and the Fermi Paradox. International Journal of Astrobiology 12(2), 144–148.
- Drake FD (1961) Discussion of Space Science Board, National Academy of Sciences Conference on Extraterrestrial Intelligent Life. Green Bank, West Virginia.
- Drexler KE (1987) Engines of Creation. New York: Anchor, pp. 172-173.
- Duncan RC (1991) The life-expectancy of industrial civilization, in 'System Dynamics '91: Proceedings of the 1991 International System Dynamics Conference, Bangkok, Thailand, August 27 through 30, 1991', pp. 173–181.
- Dyson FJ (1960) Search for artificial stellar sources of infrared radiation. Science 131(3414), 1667–1668. (http://science.sciencemag.org/content/131/3414/1667)

- European Hospital and Healthcare Federation (2009) Hospitals in the 27 Member States of the European Union. Paris: Dexia Editions, p. 47. (http://www.hope.be/wpcontent/uploads/2015/11/792009OTHERHospitals-in-27-Member-States-of-the-European-Union-eng.pdf)
- Farnsworth KD, Nelson J and Gershenson C (2013) Living is information processing: from molecules to global systems. *Acta Biotheoretica* **61**(2), 203–222. (https://doi.org/10.1007/s10441-013-9179-3)
- Forgan D (2009) A numerical testbed for hypotheses of extraterrestrial life and intelligence. *International Journal of Astrobiology* 8(2), 121–131.
- Frank A and Sullivan WT (2016) A new empirical constraint on the prevalence of technological species in the universe. *Astrobiology* **16**(5), 359–362.
- Hanson R (n.d.) The great filter are we almost past it?, Downloaded from: (http://mason.gmu.edu/~rhanson/greatfilter.html.), Accessed June 12, 2017.
- Herfst S, Schrauwen EJA, Linster M, Chutinimitkul S, de Wit E, Munster VJ, Sorrell EM, Bestebroer TM, Burke DF, Smith DJ, Rimmelzwaan GF, Osterhaus ADME and Fouchier RAM (2012) Airborne transmission of influenza A/H5N1 virus between ferrets. *Science* 336(6088), 1534–1541. (http://science.sciencemag.org/content/336/6088/1534)
- Imai M, Watanabe T, Hatta M, Das SC, Ozawa M, Shinya K, Zhong G, Hanson A, Katsura H, Watanabe S, Li C, Kawakami E, Yamada S, Kiso M, Suzuki Y, Maher EA, Neumann G and Kawaoka Y (2012) Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature* 486 (7403), 420–428. (http://dx.doi.org/10.1038/nature10831)
- Jackson RJ, Ramsay AJ, Christensen CD, Beaton S, Hall DF and Ramshaw IA (2001) Expression of mouse interleukin-4 by a recombinant ectromelia virus suppresses cytolytic lymphocyte responses and overcomes genetic resistance to mousepox. *Journal of Virology* 75(3), 1205–1210. (http://jvi.asm.org/content/75/3/1205.abstract)
- Johansson F et al. (2013) mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 0.18), December 2013. (http://mpmath.org)
- Jones EM (1981) Discrete calculations of interstellar migration and settlement. Icarus 46(3), 328–336. (http://www.sciencedirect.com/science/article/pii/ 0019103581901366)
- Kardashev NS (1964) Transmission of Information by Extraterrestrial Civilizations. Soviet Astronomy, 8(2), 217–221.
- **Kompanichenko VN** (2000) Average lifetime of an intelligent civilization estimated on its global cycle, *in* 'Bioastronomy 99: A New Era in the Search for Life', Vol. 213 of Astronomical Society of the Pacific Conference Series, pp. 437–440.
- Maccone C (2010) The statistical Drake Equation. Acta Astronautica 67, 1366–1383
- McNeill WH (1976) Plagues and Peoples. Garden City, NY: Anchor. p. 94, 113ff, 152, 180, 186.
- National Cancer Institute (2018) Cancer moonshot. https://www.cancer.gov/research/key-initiatives/moonshot-cancer-initiative.
- Oliver BM and Billingham J (1971) Life in the universe, in *Project Cyclops: A Design Study of a System for Detecting Extraterrestrial Intelligent Life*, Stanford/NASA/Ames Research Center. NASA report CR 114445, pp. 3–28. (https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19730010095.pdf)
- Organisation for Economic Cooperation and Development European Union (2016) *Health at a Glance: Europe 2016*, Paris: OECD Publishing, p. 193. (http://dx.doi.org/10.1787/9789264265592-en)
- Perelson AS, Rong L and Hayden FG (2012) Combination antiviral therapy for influenza: predictions from modeling of human infections. *The Journal of Infectious Diseases* 205(11), 1642–1645. (http://dx.doi.org/10.1093/infdis/jis265)
- Phillips JA (1978) Mushroom: The Story of the A-bomb Kid. New York: Morrow.
 President's Council of Advisors on Science and Technology (2016) Action needed to protect against biological attack. https://obamawhitehouse.archives.gov/blog/2016/11/15/pcast-letter-president-action-needed-protect-against-biological-attack.
- Rozo M and Gronvall GK (2015) The reemergent 1977 H1N1 strain and the gain-of-function debate. mBio 6(4), e01013–15.
- Rubin CT (2001) L factor: hope and fear in the search for extraterrestrial intelligence. *Proc. SPIE* 4273, 230–239. (http://dx.doi.org/10.1117/12.435379)
- Sandberg A, Armstrong S and Cirkovic MM (2017) That is not dead which can eternal lie: the aestivation hypothesis for resolving Fermi's paradox. arXiv:1705.03394. https://arxiv.org/abs/1705.03394

Schenkel P (1999) The nature of ETI, its longevity and likely interest in mankind - the human analogy re-examined. *Journal of the British Interplanetary Society* **52**, 13–18.

Shklovskii IS and Sagan C (1966) Intelligent Life in the Universe. New York: Delta, pp. 409–418.

Suskind R (2006) *The One Percent Doctrine*. New York: Simon and Schuster, p. 62.

Tipler FJ (1980) Extraterrestrial intelligent beings do not exist. *Q J R Astro Soc* **21**, 267–281.

Turnbull MC and Tarter JC (2003) Target selection for SETI. II. Tycho-2 dwarfs, old open clusters, and the nearest 100 stars. Astrophysical Journal Supplement Series 149, 423–436.

Vakoch DA and Dowd MF (2015) The Drake Equation: Estimating the Prevalence of Extraterrestrial Life through the Ages. New York: Cambridge University Press. von Hoerner S (1961) The search for signals from other civilizations. Science 134, 1839–1843.

Watson C, Sell TK, Watson M, Rivers C, Hurtado C, Shearer MP, Geleta A and Inglesby T (2018) *Technologies to Address Global Catastrophic Biological Risks*. Baltimore: Johns Hopkins University Bloomberg School of Public Health/Center for Health Security, p. 8. (http://www.centerfor healthsecurity.org/our-work/pubsarchive/pubs-pdfs/2018/181009-gcbr-tech-report.pdf)

Webb S (2015) *If the Universe Is Teeming with Aliens ... Where is Everybody?*, 2nd Edn. Cham, Switzerland: Springer.

Wertheim JO (2010) The re-emergence of H1N1 influenza virus in 1977: A cautionary tale for estimating divergence times using biologically unrealistic sampling dates. PLOS ONE 5(6), 1–4. (https://doi.org/10.1371/journal.pone.0011184)

Zimmer SM and Burke DS (2009) Historical perspective – emergence of influenza A (H1N1) viruses. New England Journal of Medicine 361(3), 279–285. PMID: 19564632. (http://dx.doi.org/10.1056/NEJMra0904322)

Mathematical Appendix

Math 1: ln(1 - P) = -P *as* P → 0

To solve

$$f(P) = \frac{ln(1-P)}{P}$$

at P = 0, we observe that f(0) evaluates to 0/0, making the expression indeterminate. However, it also means L'Hôpital's rule applies in the second step below:

$$\begin{split} \lim_{P \to 0} f(P) &= \lim_{P \to 0} \frac{\ln(1-P)}{P} \\ &= \frac{\lim_{P \to 0} (d[\ln(1-P)])/(dP)}{\lim_{P \to 0} (d[P])/(dP)} \\ &= \frac{\lim_{P \to 0} ((-1)/(1-P))}{\lim_{P \to 0} (1)} \\ &= \frac{-1}{1} \end{split}$$

Hence:

$$\lim_{P \to 0} \frac{\ln(1-P)}{P} = -1$$

So, when $P \rightarrow 0$ we can use:

$$ln(1-P) = -P \tag{A1}$$

For our purposes this approximation is excellent, viz. ln(1-0.1) = -0.105 and ln(1-0.001) = -0.0010005.

Math 2: LD_{50} as $P \rightarrow 0$

We start with equation (3) defining LD_{50} , then simultaneously take the limit and substitute equation (A1) into it:

$$LD_{50} = \frac{\ln(1 - 0.50)}{E \times \ln(1 - P)}$$

$$\lim_{P \to 0} LD_{50} = \frac{\ln(1 - 0.50)}{E \times (-P)}$$

$$= \frac{\ln 2}{E \times P}$$
(A2)

$$\approx \frac{0.7}{E \times P} \tag{A3}$$

Thinking solely in terms of exponents:

$$LD_{50} \approx 0.7 \times 10^{-(log_{10}E + log_{10}P)}$$

Math 3: N(w) - Exact

Recall from equation (1) that C(y) is the fraction of civilizations still communicating y years after their birth. Here, however, the notion of time changes a bit

First, define:

B(y) = number of new civilizations <u>b</u>orn in year y

N(y) = number of communicating civilizations existing in year y

Next, assume we are interested in a window of time in the galaxy's history running from year 0 to year w, where no civilizations were present at y = 0. We want to know the number of communicating civilizations that exist at the end of the window, i.e. at time w.

To be considered alive at year w, any civilization born in some year y will have to communicate for w-y more years. Thus:

$$N(w) = B(0) C(w) + B(1) C(w-1) + B(2) C(w-2) + \dots + B(w) C(0)$$

Assuming B(y) is a constant (having units: civ year⁻¹):

$$N(w) = B \sum_{y=0}^{w} C(y)$$
 (A4)

We can replace summation with integration:

$$N(w) = B \int_0^w C(y) \, \mathrm{d}y \tag{A5}$$

To solve for N(w), assuming all civilizations have the same E and P, we define:

$$S = (1 - P)^E \tag{A6}$$

Substituting the above into equation (1) yields:

$$C(y) = S^{y} \tag{A7}$$

Then substituting equation (A7) into equation (A5):

$$N(w) = B \int_{0}^{w} S^{y} dy$$
$$= B \frac{S^{y}}{\ln S} \Big|_{0}^{w}$$
$$= B \left(\frac{S^{w}}{\ln S} - \frac{S^{0}}{\ln S} \right)$$

This yields the exact form of N(w):

$$N(w) = B \frac{S^{w} - 1}{\ln S}$$
 where $S = (1 - P)^{E}$ (A8)

Math 4: N(w) - As P \rightarrow 0 and w $\rightarrow \infty$

In many scenarios for N(w), $P \to 0$ and/or $w \to \infty$. We here derive an approximation for such conditions.

First, expand the exact definition of N(w) in equation (A8):

$$N(w) = B \frac{((1-P)^E)^w - 1}{\ln((1-P)^E)}$$
$$= B \frac{(1-P)^{(Ew)} - 1}{E \ln(1-P)}$$

Now substitute with the results of equation (A1), namely ln(1-P) = -P when P is small and, consequently, $(1-P) = e^{-P}$:

$$\lim_{P \to 0} N(w) = B \frac{e^{-PEw} - 1}{E(-P)}$$
$$= B \frac{1 - e^{-PEw}}{EP}$$

As w becomes large, $e^{-PEw} \rightarrow 0$. Thus:

$$\lim_{P \to 0} N(w) = \frac{B}{EP} \tag{A9}$$

Using Fig. 3, which was calculated using the exact form of N(w) in equation (A8), we observe the approximate value-range of w for which the limit of equation (A9) holds:

$$N(w) \approx \frac{B}{EP} \left[\text{for } w \ge \frac{10}{EP} \text{ and small } P \right]$$
 (A10)

Math 5: $\overline{L(w)}$

As many others have noted, the Drake equation can be reduced to a two-parameter form:

$$N = B \times L \tag{A11}$$

where N is the number of communicating civilizations, B is the birth rate of communicating civilizations and L is the mean lifetime of all birthed civilizations.

Applying this to our approach of examining time windows having constant *B*, we can rewrite equation (A11) as:

$$N(w) = B \times \overline{L(w)} \tag{A12}$$

where $\overline{L(w)}$ is the mean lifetime of a civilization born during the time window that extends from 0 to w.

Rearranging equation (A12) and then substituting from equation (A8) yields:

$$\overline{L(w)} = \frac{1}{B}N(w)$$

$$= \frac{1}{B}B\frac{S^{w} - 1}{\ln S}$$

$$= \frac{S^{w} - 1}{\ln S}$$

$$= \frac{(1 - P)^{Ew} - 1}{\ln (1 - P)^{E}}$$
(A13)

To derive a simple approximation for $\overline{L(w)}$, recall from equation (A10) that $N(w) \approx B/(EP)$. It is immediately apparent from equation (A12) that:

$$\overline{L(w)} \approx \frac{1}{EP} \left[\text{for } w \ge \frac{10}{EP} \text{ and small } P \right]$$
 (A14)

Finally, the ratio of equations (A2) to (A14) is noteworthy:

$$LD_{50} / \overline{L(w)} \approx \ln 2 \approx 0.7$$
 [for $w \ge \frac{10}{EP}$ and small P]

Math 6: Maximum $\overline{L(w)}$

To find the *w* where $\overline{L(w)}$ is maximal, we set the derivative of the definition of $\overline{L(w)}$ (from equation (A13)) to zero:

$$0 = \frac{d[(S^w - 1)/(\ln S)]}{dw}$$
$$= \frac{1}{\ln S} S^w \ln S$$
$$= S^w$$

Given 0 < S < 1, then $S^w = 0$ at $w = \infty$. So, using equations (A13) and (A1):

$$\overline{L(w)}_{\text{max}} = \overline{L(\infty)} = \frac{S^{\infty} - 1}{\ln S}$$

$$= \frac{-1}{E \ln(1 - P)}$$

$$= \frac{1}{EP} \text{ as } P \to 0$$

Seeking to show $\overline{L(w)}_{\text{max}} < 1/(EP)$ for all P, we begin by observing:

$$x - \frac{1}{2}x^2 + \frac{1}{3}x^3 \dots < x$$
 [for $x < 0$]

The Mercator series is:

$$x - \frac{1}{2}x^2 + \frac{1}{3}x^3 \dots = \ln(1+x)$$
 [for $-1 < x \le 1$]

We combine the two preceding formulae into an inequality, then set x = -P:

$$ln (1+x) < x$$
 [for $-1 < x < 0$]
 $ln (1-P) < -P$ [for $-1 < -P < 0$]

Substituting back into the definitions of $\overline{L(w)}_{max}$ gives, for 0 < P < 1:

$$\overline{L(w)}_{\text{max}} = \overline{L(\infty)} = \frac{-1}{E \ln(1-P)} = \left(\lim_{P \to 0} \frac{1}{EP}\right) < \frac{1}{EP}$$
(A15)

Math 7: Model of E Growing Over Time

We wish to model *E* growing over time and apply the model to time spans that do not cause exponential explosion.

First, recall equation (1), the basis for the baseline 'Constant-E' model:

$$C(y) = (1 - P)^{(Ey)} \tag{1}$$

The exponential term Ey has units entity-years and signifies the total exposure of the civilization to destruction events. Renaming E to E_0 to reinforce its constant nature in the Constant-E model, we can write:

Exposure[Constant Model]_{$$\nu$$} = $E_0 y$ (A16)

Similarly, we can calculate exposure for a model in which E grows with time. Equation (6) defined E_y as growing from an initial value of E_0 at an annual rate of r over a period of y years, with the growth continuously compounded:

$$E_{\nu} = E_0 e^{ry} \tag{6}$$

In this 'Growing-E' model, the civilization's exposure to destruction events is:

Exposure[GrowthModel]_y =
$$\int_{t=0}^{y} E_0 e^{rt} dt$$

= $E_0 \left(\frac{e^{rt}}{r} \Big|_{t=0}^{y} \right)$
= $E_0 \frac{e^{ry} - 1}{r}$ (A17)

We can equate the two exposures from the right-hand-sides of equations (A16) and (A17), taking care to distinguish the two different y:

$$E_0 y_1 = E_0 \frac{e^{ry_2} - 1}{r}$$

This equation says that a civilization's destruction-exposure after y_1 years as calculated by the constant model, equals the exposure after y_2 years as calculated by the growth model.

Continuing, we can cancel the E_0 terms and express y_2 in terms of y_1 :

$$y_1 = \frac{e^{ry_2} - 1}{r}$$

$$y_2 = \frac{\ln(1 + ry_1)}{r} \tag{A18}$$

Analytically, equation (A18) provides a shortcut for converting results from the constant-E model to the growing-E model. For example, we can define the LD_{50} for the growing-E model as:

$$LD_{50}[G] = \frac{\ln(1 + rLD_{50}[C])}{r}$$
(A19)

where the [G] and [C] indicate the growing-E model and constant-E model, respectively. See Fig. 4.

Although equation (A18) does not have an explicit exponential term, it must still be applied carefully because it implicitly assumes that the number of entities can grow exponentially without limit, per equation (6). Alternative growing-*E* models may be derived, e.g. using linear growth.