Response to reviewers.

Dear reviewers,

Thank you so much for your helpful reviews. We've given every effort to address the points you raised in as direct a way as possible, and this has changed the form of the manuscript from the first version you read. Some sections have been eliminated, a new one has been created, and we've made global changes for the sake of clarity. The largest revisions are the following:

1) We've given a full generalization of the Lexis diagram expressed mathematically, such that the hexad identity we focus on is a particular case. This development gives a more solid basis to our proposal, and it also gives a common language to facilitate comparison with other temporal designs, such as the Lexis marriage identity, and a more complex design given in Brinks et al (2014) (reference given by reviewer 1). This is the most important single change to the manuscript.

2) Certain sections have been outright eliminated, including “Relationships between the diagrams”, “A geometrical analogy”, “The informative triads”, “The extension of time axes”, and “Related concepts”. Most of these sections contained non-essential observations, and we agree that they may have reduced readability.

3) “Discussion” now is “Conclusions”, and it is structured, briefer, and to-the-point. We think this it is OK to omit a discussion section since concepts are discussed throughout.

4) Throughout the manuscript we've tried to enhance appeal with plain and/or rigorous language.

We address your comments point-by-point:

> Reviewer #1: This manuscript forms a beginning of a possibly interesting generalization of the Lexis diagram, but is rather incomplete regarding
> *Precise mathematical statements
> *Satisfactory explanations of applications
> *Reference to many important current papers

All points now accounted for. Although we've not reproduced formulas in the “Application” section, we did try to do a better job of explaining it while not lengthening it. In a future manuscript we aim to describe a potluck of diverse applications.

> Lexis's three-dimensional diagram is discussed but never shown. The diagram itself was reproduced by Keiding (2006) and a very informative graph of the idea was published by Brinks et al. (2014). Jewell (2016) surveyed this area. The present authors claim that Lexis's presentation includes a similar idea as their own, but they never state precisely where they see this equivalence. Actually Lexis essentially looked forward in time, just as in the illness-death model, as carefully explained by Keiding and by Brinks et al. in the above references.

The Lexis and Brinks setups are now efficiently discussed in Section 4.2, examples 2 and 4, respectively. Thank you for the Brinks reference, as our reading of that paper leads us to conclude that the simulation model itself has four point time measures (events), (birth, period, diagnosis, death), and these imply a total of 6 possible duration time measures, which we show can be represented in a hypothetical four-dimensional diagram. We explain how their 3d Figure is based on a subset of the time measures implied by their simulation model, and we think that that was a perfectly reasonable choice since a 4d diagram would be hard to usefully represent.

> On the contrary the authors look at reverse time, for this literature see Chan and Wang (2010), Dempsey and McCullagh (2016) and the general survey by Kurland et al. (2009).

We found these references helpful in various places in the manuscript, thank you. We hadn't seen the term

“reverse time” before.

> The equilateral ('isotropic') diagram was discussed by Lexis (1875) and Lewin (1876) but not by Zeuner.

Thank you for the correction. That error was embarrassing.

> Applications are sketched far too loosely. Examples: three lines above Section 3.3 (p.11); three lines above Section 3.3.3 (p. 15) - here the key point should be that experience seems to indicate that sometimes it is more informative to count backwards from terminal events, see the above references; seven lines before Section 4, p. 18.

We've given a few more specific examples in the 2d diagram sections. We've not developed any of these examples in more detail, as this would lengthen the paper, which we needed to shorten. Some of the suggested references were useful in these sections.

> The big formal contribution seems to be the tetrahedron in Section 4. But this is too poorly explained for the non-mathematician and proper mathematical definition and interpretation are lacking.

We hope that this issue is now resolved, given the more general event-duration framework proposed, and the description and examples of graph construction provided in new Section 4.

> The various references to stationary populations may well be interesting for further study - but they are really not at the core of the diagrams.

We've removed the stationary population discussion for the most part.

Selected references

Brinks R, Landwehr S, Fischer-Betz R, Schneider M, Giani G (2014) Lexis Diagram and Illness-Death Model: Simulating Populations in Chronic Disease Epidemiology. PLoS ONE 9(9): e106043.

Chan,K.C.G.. & Wang, M.C. (2010). BACKWARD ESTIMATION OF STOCHASTIC PROCESSES WITH FAILURE EVENTS AS TIME ORIGINS. The Annals of Applied Statistics 4, 1602-1620

Dempsey, W. and McCullagh, P. (2016). Survival models and health sequences. https://arxiv.org/pdf/1301.

Jewell, N.P. Natural History of Diseases: Statistical Designs and Issues. CLINICAL PHARMACOLOGY & THERAPEUTICS | VOLUME 100 NUMBER 4 | OCTOBER 2016

Keiding N (2006) Event history analysis and the cross-section. Statistics in Medicine 25: 2343-2364.

B.F. Kurland, L.L. Johnson, B.L. Egleston & P.H. Diehr (2009). Longitudinal data with follow-up truncated by death: Match the analysis method to research aims. Statistical Science 24, 211-222.

Thanks for passing these on. Each has been considered, and most are now accounted for. We didn't go further than this, because we think there are enough examples mentioned and because we don't think about this material (usually) as a matter of statistical design. Because we're not statisticians. But it's nice to know that some of these configurations have been considered, and we'll try to keep tabs on developments in this area.

Reviewer 2:

> Dear Authors, Dear Editor,
>
> This is an interesting and valuable contribution that I enjoyed reading and that should be of interest to those interested in basic demographic measures and mathematical demography. However, the paper is also rather complex and long, so I urge the authors to go carefully again through the document in an attempt to find any possibility to improve readability to the readers. Readability might also be improved by providing more examples and more intuitive interpretations throughout the manuscript. In addition I would like to ask the authors to have a look at the discussion, as the structure and the arguments should be improved in this section. Despite these shortcomings I feel that the paper is in a rather good state. Therefore, most of the comments that I would like to offer in the following are rather minor or rather general calls for more clarity.

Thank you. We've taken these suggestions seriously. On the whole the paper is shorter, and we've tried to offer plain language explanations of things where appropriate, and rigorous clarifying language where helpful. The new mathematical section 4 (the event-duration generalization) may appear to reduce the legibility of the manuscript due to terse mathematical language, and we try to offset this with a set of four very brief examples, and the use of some visuals. This addition was necessary in order to give a firm grounding to our diagram and to make it comparable with other temporal designs. As mentioned, we've done an almost complete redraft of the discussion.

> p.2 "lifespan (L) or age-at-death"
>
> I was wondering whether the term "age at death" should be used throughout the manuscript, it feels more straightforward than lifespan. Lifespan has number of different meanings, sometimes it is referred to the maximum age at death that a species can obtain, sometimes it might refer to the average length of life for certain species. It seems more rarely used to define the length of life for an individual, therefore I would prefer age at death.

This is a valid observation. We've kept lifespan L, but tried to be clearer that this is *individual* lifespan or age at death in a couple places.

> p.2 "This popular representation does not account for remaining years of life and other related time indices that may be of interest to researchers and policy makers."
>
> This sentence is both used in the Abstract and the introduction but seems too generic and not a good justification for your work. Who are these "related indices", why would we care about including others than the ones we use (APC). What are the common characteristics of those we would be interested in or that you study? There could be a number of ways to improve this introduction sentence into something more meaningful.

We've changed the sentence to “This relationship does not account for remaining years of life, total length of life, or time of death, whose use in demographic research is both underrepresented and incompletely situated.” We hope that motivations for doing so are sufficiently reinforced by the mention or suggestion of applications (all of which are brief), but in general we try to make a case for *pattern detection* as a proximate goal of this and other potential new diagrams.

> p.5 "terminal state"
>
> "absorbing state" would fit better?

Nice catch.

> p.5 "It is the measure we all want to know, the thing we approximate with remaining life expectancy."
>
> The sentence could be rewritten, too colloquial for my taste.

The sentence now reads “Thanatological age is remaining time until death, the information approximated with life expectancy.”

> p.8 "The isotropic mapping."
>
> In this paragraph it remains unclear why it would matter to have a different projection, except for the fact that you have simply

rotations (which usefuleness might be unclear (at this point). In the Lexis diagram the derived time measure is longer but it doesn't matter because nobody? is using the Lexis diagram to measure the length of life with a ruler. If there are any other reservations to share (from Perozzo or Zeuner for example), it would be good if the authors could mention them here. Or provide example why it would matter for some of the other triads in question already.

We've removed all 2d isotropic diagrams from the manuscript to reduce clutter, but that paragraph remains and includes a sentence “The primary justification for isotropic demographic surfaces comes from a data visualization perspective, where it may be hypothesized that the viewer’s ability to compare slopes is hindered if time coordinates are not on the same scale.”

P12. Footnote: Some prefer the term Lexis surface, while others prefer to call them contour maps, heatmaps, or stereograms.

I would suggest removing this footnote. Lexis surfaces is exactly what you are referring to, all other terms are much more general terms.

Done.

p.18 "A space derived by extending any of the triad identities into its lacking dimension implies each of the other triad identities, making a total of six time indices. In essence, the four triad identities may be thought of as the four faces of a tetrahedron."

You lost me here. Any way to write this part more clearly?

Sentence removed (and some others too here). This section is redrafted to be shorter, and to follow from the more general framework given in new section 4.

p.21 " of the the above-described tetrahedron"
typo

Fixed.

p.21 / p.22 Section 4.3
I think it might be good to provide the aim of this part to the reader at the beginning of it. You lost me here as well as it is difficult to read.

Section removed.

p.26 "and reflection that was based this coordinate system,"
Typo?

The meaning is now captured in “... it is based on a relatively minor tweak to standard methodology, itself inspired by viewing data under the conditions enabled by the demographic time framework and adjusting standard demographic methods to capture the direction of temporal variation in data.”

p.26 The last paragraph on the page is not needed, I suggest to shorten it to "Let us take the example of self-reported health (SRH)."

Done.

p. 27. "Figure 7 displays a series of TAL surface plots, each referring to a different quinquennial birth cohort (1905-1909, etc)"

To help the reader, I suggest to add a sentence or two here that clearly state what kind of results we can get from these.

Done, thanks for pointing this out. We added a sentence saying “We opt to view data on TAL surfaces because these allow us to judge the shape of prevalence over the lifecourse, specifically to show how SRH prevalence variation is summarized more efficiently as a time-to-death pattern than as an age pattern.”

p.27 ", in this case over thanatological age."

I suggest dropping this half-sentence as it is repeated two-sentences later (but there it seems to fit better)

Done. Thanks.

> p.27 "Contour lines in the surfaces indicate the primary direction of variation, in this case over thanatological age. Downward diagonals indicate lifespans, which the reader may also think of as very specifc birth-death cohorts. These are the diagonals along which lifelines may be imagined, as suggested in Figures 3 and 6."
>
> I think the reader would benefit with a more clear (reminder) of the interpretation here, especially as it might be that some readers may jump to the example directly.

Explanation now given.

> p.28. If we apply this trajectory to the synthetic stationary population of each year from 1980 and 2010 (HMD 2016), we can calculate the resulting healthy and unhealthy life expectancies, and compare these with the expectancies that we would have projected assuming the 1980 Sullivan curve.
>
> The remains rather unclear, please improve the description of the approach.

I'm not sure if we've succeeded here, but we do not wish to include formulas for this in this very manuscript. This now reads:

> “Let us take the population of US males aged 60 and older, and assume that mean time-to-death trajectory derived from the Fig. 7 surfaces is valid for them. We apply this trajectory to the synthetic stationary population of each year from 1980 and 2010 (HMD 2016) following the formulas in Van Raalte & Riffe (2016). We then calculate the resulting healthy and unhealthy life expectancies, and compare these with expectancies calculated using the standard Sullivtan method and assuming the 1980 chronological age pattern of poor SRH.”

If this is deemed an insufficient description, then we think the next best option would be to cut back the application to just pattern detection and say somewhere “these patterns would be invisible without such surfaces.” That would be shorter, but we think readers would find the full summarized chain of inquiry more compelling because it concludes with an actual lesson. It's still a very particular application, and we do think that the framework given is much more broadly applicable.

> p.31 Start the discussion with a short summary of the paper. As said before, this section feels rather unstructured and unclear.

It's now been reformatted to a more-structured and briefer “Conclusions”

> Finally, I would like to point out that the citations and references likely suggest the identity of (one) the author(s). I think that the author(s) should have increased their efforts to provide a fully blinded manuscript.

Sorry for that. Then this review was single-blind, which is fine for us on the blind end of it.

Both reviewers were constructive and helpful, and both reviews prompted major changes to the manuscript. We think the paper is much stronger now, we're grateful for your contributions, and this will be duly acknowledged.