Thanks to the editor and authors for the opportunity to review this interesting paper. This study revisits a previously dated zircon-bearing carbonate bed, the Likhall bed, which has important implications for the Ordovician time scale but has yielded inconsistent dates between studies. The authors attribute these differences to both

- methodological issues (mainly incomplete chemical abrasion by previous studies, as well systematic differences in tracer calibration between studies using an ET spike and a study using an in-house spike) and
- interpretive differences relating to how one determines a single age for a bed that contains multiple dated zircon grains.

The authors grapple with these problems and put forward a new weighted mean age estimate for the bed. The study is well-written, concise, and clear. I support publication of the study and make the following (hopefully constructive) suggestions.

This bed was previously dated by two studies, Lindskog et al. 2017 and Liao et al. 2020, and ages from all three studies are shown in Figure 2.

- This figure shows all the ages with their X uncertainty, but I believe they should all be shown with their Y uncertainty (including tracer calibration uncertainty), as the Liao data were determined using an in-house spike and the Lindskog/Paul data were generated using an ET spike. If the authors don't wish to or can't do this, then I suggest caveating this in the caption.

Line 95: Can the authors provide a citation supporting the assumption of a magma initial Th/U ratio of 3.5 +/- 1?

Figure 1:

- I suggest recoloring this figure so that the cut-off value of 50 is at an inflection point in the colorscale and/or indicate on the colorscale where 50 is. It is hard to look at this and immediately understand that some of the green grains that overlap with the concordia curve are actually below 50 and should be left out.
- It's also somewhat confusing that the grains "selected" for analysis here are not the same as the grains selected in Table S1, column "interpretation strategy i."
    - I suggest annotating this column in Table S1 to indicate the three grains that are selected for analysis using the more restrictive Pb*/Pbc screen.
    - I suggest adjusting the figure so that all grains that are "selected" using "interpretation strategy i" have a black dashed ellipse and those that are "selected" using the more strict Pb*/Pbc screen have a black ellipse (as they do now).
- Isn't a part of the selection criteria also that the ages are within the youngest cluster (line 305)? This should be stated in the caption.

Line 125: I think the geological setting should be given greater priority within the manuscript and encourage the authors to include a figure showing the field context of the bed sampled. I understand that this has been described elsewhere, but such a figure would increase the value of this manuscript to the reader. Similarly, images of the zircon grains have been previously published elsewhere, but including some images in this manuscript or in the supplement would be helpful.

The authors first consider differences in the chemical abrasion techniques used between studies, and argue that incomplete chemical abrasion of inclusions and metamict zones within zircons resulted in inaccurate ages for the previous studies. Previous studies used lower temperatures for their chemical abrasion procedures than the current study. The main data the authors use to support this idea is shown in Figure 4, which plots Pbc vs Pb* for each of the three studies.

- This is more of an aesthetic suggestion, but I suggest making all the plots the same size and aligning their left y-axis.
- I also think that plotting all the data with the same axes or on the same plot would help the authors make their point that their approach has resulted in much smaller Pbc measurements and presumably more precise and accurate dates.
- Because chemical abrasion is meant to remove zones of zircon affected by Pb loss, it's easy to follow this argument for the Liao data, which the authors argue is too young because of incomplete chemical abrasion, but harder to follow it for the Lindskog data, which is too old (and the authors discuss this). They suggest that the "too old" ages come from inaccurate and imprecise blank corrections.
    - Line 231: "If we assume that larger zircons contain more Pb* and a larger volume of Pbc-bearing inclusions…" I'm happy to agree that larger zircons would have more Pb* and Pbc, but I'm missing a step in the logic of the sentence here. I think it's that larger zircons with more Pbc would be more impacted by an erroneous blank correction, but perhaps it can be spelled out a bit more clearly.

The authors consider several different interpretation strategies. They identify five ways that one might determine the age of this bed:

1. Weighted mean of a subset of data
2. Youngest cluster of overlapping ages at 2-sigma
3. Use entire range of concordant zircon analyses as autocrystic growth in magma chamber
4. Use the youngest concordant grain as the best proxy for the timing of eruption
5. A Bayesian approach as suggested by Keller et al. 2018

- I would rephrase 5 to something like "A Bayesian approach such as those suggested by Keller et al. 2018 or Traylor et al. 2021."
    - The authors immediately state that #5 is outside the scope of this study, but I don't agree. The approach used by Keller et al. 2018 is relatively easy to implement and a publicly-available Jupiter Notebook provides help at https://github.com/brenhinkeller/Chron.jl?tab=readme-ov-file. I suggest that it be included and discussed; if the authors have a major disagreement with this approach, they should show why.
    - Traylor et al. refers to modifiedBChron, which would consider all analyses as part of a summed probability density distribution function and then use stratigraphic superposition in a Bayesian model to determine an age that is supposed to be a better representative of geological uncertainty. Using only the new data produced by this study, of course this can't be applied as we don't have the benefit of stratigraphic superposition, but if, as the authors state, there are other dated bentonites in the section with clear stratigraphic relationships, why not give it a try, and see how it compares? modifiedBChron is also fairly easy to implement. I

suspect the outcome will mostly highlight the importance of getting good age constraints elsewhere in the section, but this would be worth highlighting too.

The Bayesian approaches introduced by Traylor and Keller are exciting new developments that help us understand what the meaning of a dated ash bed means in stratigraphic context. This paper's ambition is to make the point that the Ordovician timescale must be reconsidered, and it certainly provides evidence in that direction, but at present, this paper misses an opportunity to grapple with these emerging approaches and their implications for the timescale.

The authors spend a lot of time discussing using Pb*/Pbc as a screening metric and note that it significantly improves accuracy while coming at a significant cost of greatly reducing the number of grains that are viable. Can they give non-zircon geochronologists some sense of how a strict requirement of Pb*/Pbc > 50, for example, would affect the universe of published CA-ID-TIMS zircon data? Would it knock out >50% of published zircon grains, as it does here? Given the expense and time-consuming nature of CA-ID-TIMS analysis, how do they recommend other workers grapple with this?

Line 299: "leaves only 8 of 22"—yes and worth stating explicitly that these 8 are not all within the youngest age cluster.

The authors describe some of the changes to the Ordovician timescale that their new age requires.

- It would be helpful for the authors to visualize these changes in a "before" and "after" figure.
- The authors also note that the timing of L-chondrite breakup should be revised to c. 467.1 Ma, but the reader isn't given enough context about why this date should be assigned to the breakup, (which isn't the age they give the Likhall bed).
- As the authors motivated this study by mentioning the controversial hypothesis linking meteorite bombardment with Ordovician biodiversification in the introduction, they should return to this hypothesis and discuss the alignment or un-alignment of these two events in light of their new results.

Finally, the references:

- Liao et al. 2020 is a key study for this paper and the source of much of the legacy data discussed, but it does not appear in the reference list. This must be addressed before final publication.
- I suggest reformatting the Schmitz et al. citation in the references to superscript 205, 235, 238, etc. Same thing for the von Quadt et al. citation and 10^13.
- On the topic of CA-ID-TIMS dating of cryptotephra from carbonates: The authors might be interested to know about a similar study focused on retrieving volcanic zircon from carbonate rock: Finzel and Rosenblume 2021, *Geology*. I think referencing this study would help make the authors' point that determining best interpretive practices for carbonate-derived crypotephra ashes is a question with applications beyond the single bed considered here.