Reviewer 1		
Line	**Reviewer comment**	**Response**
		We thank the reviewer for their comments. We emphasise that our paper proposes a way to estimate risk to life in the absence of any information apart from regional-scale morphometric analysis such as the Melton ratio. Many of the reviewer's comments regarding our assumptions are correct—but note that our aim is to demonstrate the possibility of risk to life, where no risk is perceived by communities or their decision-makers. To do this easily and cheaply on a regional scale, we need to make “precautionary but realistic" assumptions. These assumptions therefore will err on the side of caution. If our analysis gains the attention of communities and their decision-makers, we are able to investigate in more detail.
	However, this key finding is currently vague, partly because it is difficult to estimate risk parameters and partly because the current manuscript has limited/unclear presentation of the specific return intervals that result in unacceptable risk. Are they intended to be global/generalizable estimates, or regionally specific? Fully characterizing hazard and risk for decision-making would require additional process-based investigation or empirical approaches that use local/regional data. I encourage the authors to describe how their model results could be used to support actionable strategies for prioritizing further risk-reduction efforts.	Our purpose is simply to demonstrate the need for risk reduction—prioritisation would be the next step. At least in NZ, but we suspect elsewhere, the risk to life and property from debris flows is often ignored by communities and their decision-makers. So even to get acknowledgement of the potential for a problem is an achievement—it is not a trivial task. We will bring this point to the fore in any revision.

Line	Reviewer comment	Response
	Significant re-framing to better reflect the conclusions. This could involve a more specific paper title, updates to the abstract, introduction, and conclusions that better reflect the contributions of their work. Specifically, I would suggest that the authors re-align the work with the primary contribution as described in the first paragraph of this review.	Agreed—see our comments above.
	Deeper grounding in real-world processes. This could be accomplished through improved basis in the literature and in consideration of both physical and social processes. Most importantly, the authors mention field evidence in the conclusions, but do not introduce study areas, present field observation, or describe data collection methodologies. These are critical for understanding the scientific contribution as well as its ability to be generalized or extrapolated to other study regions.	Our paper is a "methods" paper, and describes a straightforward method based on well-accepted principles. If one accepts the underlying principles (which are widely accepted in landside risk management) and understands that the assumed variable values are based on best available published data, it has the potential to be a valuable tool to create awareness of debris flow risks where that awareness is lacking. We do not see it as a novel contribution in scientific terms—far from it. But we believe it is a novel way to communicate debris flow risks when so many communities are oblivious or complacent. We ask that it be judged according to that criterion.
	Expanded/restructured introduction and methods section that present existing knowledge and the research question (introduction) and the author's novel approach (methods).	
	Finally, the authors describe a single, universal "window of non-recognition," which reflects the generic or best estimate parameter values for risk. However, considering that risk estimates and the window of observation relies on settlement periods, risk tolerance, and physical processes (probability of avulsion), I am not convinced that this value would apply to broad areas. Instead, I suggest that the authors present their window with a methodology framework, investigate a specific study area, and/or evaluate a wider range of parameter values for each risk parameter.	Thank you for this suggestion. However, we do ask the reviewer to reconsider this point. Our methodology can be applied to any situation where there are potential risks from debris flow hazards. However, the nature of the risks and their assessments will vary widely, depending on the factors mentioned by the reviewer (settlement periods, risk tolerance, and physical processes such as probability of avulsion). We do show how these can be included in the indicative assessment of risk to life. We have used a case study, using ARIs from two well-known life-threatening debris flow events in NZ.

Line	Reviewer comment	Response
	A few specific recommendations are described in the line comments, but in general I recommend separating background contextual information in the introduction, with specific methodology descriptions for the author's work in the methods. Currently, the methods section describes the research problem, which would be better suited in an introduction section.	Agreed.
1	Consider a more descriptive title which highlights specific analysis or finding of your work.	Agreed, we could change the title to "Identifying unrecognised risk to life from debris flows"
24	"No history of debris flows," is too vague. Do you mean geologic history? Oral history? Written history? Consider describing your study area, as settlement history is also highly variable around the world. Expanding urban areas and limited records absolutely result in low public awareness of debris flow hazards, but consider acknowledging that human settlement in New Zealand is relatively short (hence the challenge the authors' research question) densely inhabited areas in Eurasia may have hundreds of years of detailed written records, and oral histories for many indigenous peoples describe landslides or other debris flows over millennia.	A fair point. "No history of debris flows" could be rewritten as "no current knowledge about previous occurrence of debris flows". This current knowledge could be obtained from many different sources, the key thing is that the community and its decision makers are currently unaware of any hazard. Agreed, awareness of debris flows is variable around the world, but in many places (not just NZ) it is low. Sure, a long history of settlement may mean that communities have written or oral records dating back centuries or more. But we still see landslide disasters in long-settled regions such as China or South America. In many cases, rapid population increases and/or poverty have forced people to settle in areas that were previously not occupied. We do recognise NZs short human settlement history in the text (line 200 et seq.)
28-34	Please provide citations to the literature so that readers can seek additional context on debris flow processes, sediment pathways, and debris flow hazards in New Zealand.	Agreed, this would be useful, and a selection of papers can be cited.

Line	Reviewer comment	Response
Section 2.1	Section 2.1. Much of this section would be better suited to the introduction.	Agreed.
58	Table 1 does not summarize ARIs for different catchments. Please add the summary table and revise your in-text citation.	We ask the reviewer to reconsider this point. For most catchments, we do not have ARIs. We have a few reports which estimate ARIs for several well-known life-threatening debris flow catchments in NZ, that is all. Instead, we invert the problem and say that if we assume a plausible ARI (200-500 years) does this result in an unacceptable risk to life? We use a threshold of 0.001 for annual RTL, but discussion in the NZ Geotechnical Society (2023) describes annual individual fatality rates of 0.0001--our analysis can easily accommodate different choices of threshold for RTL.
62	It is true that field evidence & topographic analysis can be costly to collect and process. However, I would argue that these are the best tools for developing specific hazards understanding & precise risk estimates. Your approach for risk assessment should not replace process-based assessment, but may be useful in prioritizing communities/residences for improved outreach & risk awareness.	We certainly have no intention of supplanting the need for field work, modelling etc. Our purpose is simply to demonstrate the need for risk reduction—prioritisation would be the next step. At least in NZ, but we suspect elsewhere, the risk to life and property from debris flows is ignored by communities and their decision-makers. So even to get acknowledgement of the potential for a problem is an achievement—it is not a trivial task.
68-70	Consider “catchment gradient is associated with debris flow occurrence,” and reference other specific topographically based tools for landslide susceptibility (E.g., Montgomery et al., 1994; Dietrich et al., 2001).	Agreed, there are a broad range of methods for topographically based tools for assessing debris flow susceptibility, on both a regional and site-specific basis. A brief review of these could be included in the paper.
70-71	That identifies catchments likely to produce debris flows, which is easily calculated for many catchments over large areas, even where topographic resolution is poor or computation is limited. Also it sounds like these values have already been calculated for large areas of NZ).	The authors have completed several regional-scale investigations using Melton-R as a metric, as well as testing other methods e.g. Flow-R. However, these methods cover only limited areas of debris flow susceptible terrain in NZ.

Line	Reviewer comment	Response
96-97	I understand that you need to determine threshold risk values, but consider providing more context on how and why risk tolerance may vary, and why these values (10^{-3}-10^{-4}) are appropriate according to Taig. et al. Here or in the discussion, you may want to acknowledge that "unacceptable" risk reflects the values & tolerances of individuals and communities.	Agreed, we will revise this text to include an extended justification for choice of these threshold risk values.
Eqn 2	Are these standard abbreviations? I found them hard to follow (where does the "H" come from? Is PS:H a ratio? E stands for exposure?) Consider using simpler abbreviations or adding some description to help readers keep track of which parameter is which.	These variable names are the same or similar to the cited literature on risk to life e.g. Walker et al., 2007; Porter and Morgenstern, 2012; de Vilder et al., 2022. Since PH is the probability of a debris flow event, PS:H is the spatial (S) probability of an impact, given that the debris flow has occurred (H). We'd like to stick with this notation to ensure consistency with other published papers.
132-134/147	I'm not sure I agree with this estimate. In my experience, very few debris flows impact the entire debris flow fan. While it is challenging to estimate the probability of avulsion, are there any estimates in the literature which might describe the distribution of areas, as a proportion of total fan area, that occur during a debris flow? De Haas and others, or works by C Scheidl, DM Staley, or D Rickenmann may be useful places to look for an estimate.	It is a fair criticism. However, there is no simple, easily applied solution to this problem. We think it is more pragmatic to take an upper-bound approach here.
165-167	It may be worth adding a section in the discussion on how further investigation could be used to refine generic estimates. Sediment volume calculations, for example, could be used to improve estimates of debris flow area and deposit depths for exposure calculations. Consider expanding on this statement and adding appropriate support from the literature.	OK, but for our level of analysis we would only seek to refine estimates of variable values to a limited extent. We believe the sequence is--demonstrate the possibility of a risk to life (and/or property). Then prioritise, using regional-scale mapping of debris flow susceptibility and potential assets at risk. Would not assessment of potential debris flow severity and exposure of assets would be best focussed on studies at the site or community level?

Line	Reviewer comment	Response
183-184	These values are lower than I would expect. Can you provide more detail about the data used to make these estimates? Considering that your estimates are intended to describe worst-case scenarios where whole debris flow fans are inundated, 0.1 seems far too low.	The data were published values, so we do not have access to the data. We do not explicitly aim to model "worst-case" scenarios with complete inundation of fans. These were just credible values from the literature, and we used them to demonstrate how our model works. As we note in the text, it would be desirable to come up with variable values that were estimated or calibrated against observational data.
208	I don't agree with this assumption, and assuming the worst-case scenario conflicts with your goal of best-estimate risk calculation.	Extension of our model from individual risk to life to a "societal" risk of multiple deaths is a weakness in our model. Extending our analysis from individual risk to life to risk of multiple deaths requires knowledge of the variation amongst individuals in terms of risk variables e.g. PS:H and V. We do not have this knowledge--the best we might be able to do is assume uniform values for all individuals (but see comment re PS:H below)
Figure 1	Figure 1. What is the value shown on the y-axis?	This is a relative likelihood that that the value of the x-variable (PH) would be equal to that PH value of a random sample for the population. It is a relative likelihood, so the actual values on the axis are not so important. They could be normalised i.e. make the y-axis take values that will make the total area under the histogram equal to 1, which is saying that the total probability of the entire distribution of events is equal to a probability of 1. This would be more intuitive.
Figure 2	Figure 2. Can you provide a clearer description of the populations shown in this figure? Are these real-world catchments or monte carlo-type simulations?	See lines 230 et seq. The frequency distributions in Figs 1 and 2 are not generated from Monte-Carlo simulations. Instead, we assume a specific type of frequency distribution, the beta distribution. This has two parameters, which were chosen to correspond to a population where 95% of the population occurred within 1/500 and 1/200 for the parameter PH (annual probability of a debris flow occurrence).

Line	Reviewer comment	Response
390	No field evidence is presented here. Please add description and summary of field evidence that you use to draw conclusions.	Agreed—the way this statement is made implies that we analysed data. The comment was more of an exploratory nature—deaths from debris flows in NZ are low, therefore if we are showing that RTL may be unacceptable in debris flow catchments with settlement, we need to reconcile these two pieces of information. We will rewrite to show that our comment is intended to raise an issue, rather than resolve it.