Peer Review File

Article information: https://dx.doi.org/10.21037/apc-24-8

Reviewer A

The authors employed the Markov modeling method to estimate survival probability of stage IV pancreatic cancer. It represents a novel application of this method in this disease. However, there are significant concerns with the manuscript.

1. It is unclear whether the input data for the Markov chain analysis were representative. For instance, some studies used for parameter extraction relied on single center data. It could potentially impact the reliability of survival probability derived from these data.
Reply: yes 100% agree and this is true of any modelling used in medicine. Indeed, had the markov model shown superiority as a tool, there would need to be attempts then to address external validity. But as with any research in medicine, internal validity provides important modelling and insights as we have done here.
Changes in text: An additional paragraph has been added to explain this limitation, the value of internal validity and the threat to external validity. Importantly, we have also pointed out that by presenting the model graphically, it is very easy for others to transfer their own local data into calculation lines.

2. The Markov model of stage IV cancer depicted in Figure 5 assumed a transition stage starting from liver/lung metastases, then progressing to bone metastases, and finally to other metastases. Clinical appropriateness of this assumption is questionable. There was insufficient data to support this sequence of disease development. The order should not be determined solely by the rank of prevalence or hazard ratio.
Reply: Yes this is 100% correct which is why we outline this limitation as an assumption of the model both during model definition and as a limitation in discussion. Indeed this is an overall limitation of the Markov chain approach because there are conditions that need to be either met or assumed. This does not, however, invalidate the process. Much of what we know and understand about disease and physiology is based on modelling that must make generalizations and assumptions that are neither correct nor accurate but serve the purpose of allowing a simplified view that enhances understanding of a process. Decision tree analysis is another example of the same, where probabilities are generalized from a larger community that may or may not translate to any given sub-population. So we agree with the review note here but that is a limitation and a part of the problem (also addressed in other similar Markov papers).
Changes in text: No change as we have already recognized this as an assumption to fit Markov property satisfaction and as a limitation of the study.

3. The manuscript's presentation may not well suited for the readership of the Journal. It delved into detail of matrix calculation, which may be of limited interest to the audience.
Reply: Thanks for the comment and I do not disagree but while some readers may skip the matrix content, others will be keen to understand the simplicity of the approach.

Removing weakens the paper like it would be deleting any given section of any paper because some readers would not be interested (eg. statistical approaches sections). But importantly we informally tested the content on clinicians and found that by summarizing the maths in matrices that it actually helped them understand the approach to Markov chain approaches. So I think there is value in retaining.
Changes in text: No change because I think removing would weaken the paper, even if keeping them it means some readers skip over them.

Reviewer B

1. Authors would need to indicate the sources of the data clearly. Whether the data from a single center or multi-center, prospectively collected or retrospectively collected as well as the time frame of the data collection should be described.
Reply: I understand the comment and that is why table 1 has the citation added for each bit of data. Within word count, we can't do an analysis of search and critique methods for compiling table 1. If we were publishing a Cochrane review, we would have the space. So the question is, to what end, readers can easily see where the data comes from and determine if valid or not. And we make it clear that we are making assumptions by using data that may or may not generalize. That is the point of modelling. We have adopted data that offers complete and reliable (atleast internally in their population) data to allow our analysis to be as robust as possible. But the generalizability or external validity of those assumptions will always remain.
Changes in text: We have added a paragraph addressing this limitation in the discussion.

2. Different scenarios of metastatic disease evolvements should be considered. In the current manuscript, only liver/lung metastasis >> bone metastasis was considered. More common scenarios are liver > peritoneal metastasis, lung > peritoneal metastasis, peritoneal metastasis > liver metastasis, etc.
Reply: Yes, this is necessary on the basis of Markov assumptions and we recognize that it may not be correct (as in figure 5 makes that assumption for the purpose of the model but may not hold true). To flip figure 5 and make assumptions about the flow order does not change outcomes but does confuse the process and undermine its actual value. So it is unnecessary for the Markov component and the assumption is not made in decision tree analysis. So the modelling already addresses this.
Changes in text: No change because it has been recognized as a limitation, making the change suggested for Markov is not mathematically consistent with the very Markov properties that must be satisfied and so including would simply undermine the integrity of the model, and the order or variations to it, is already accommodated in decision tree analysis.

Editorial Comments

1. Adding more information to Table 1 as requested by Reviewer B is appropriate. There isn't a word limit for the table as far as it fits in one page.
Reply: We agree and thanks for the clarification on word limit. This is difficult because the complexity of data and analysis needed more explanation. So we have added a small

section to principal component analysis to clarify the data approach and instead of adding the data to table 1 and crowding the key information, we have added an additional table that specifically summarizes the quality of the data for both sources of table 1 and table 2. This links back to the additional text.
Changes in text: Additional explanatory text in principal components section and an additional table addressing data sources for both tables 1 and 2. We have also updated the data in tables 1-2 with recently published data not available at the time of writing the original manuscript. This has resulted in subtle adjustment to outcome results.

2. Another comment from Reviewer B is also appropriate. On one hand, it helps validate the model. On another hand, by adding scenarios more clinically relevant, the paper would fit better with the readership of this journal. New data are anticipated to be needed to address this comment and will also address Reviewer A's comment on the representativeness of the data used.
Reply: In our original Markov modelling, we needed to model based on satisfying the Markov properties. As a result, we clearly defined multiple stages that most closely resembled a uni-directional patient progression. Mathematically, it does not matter whether the proportion of people with Liver, lung or peritoneal mets – the model does not work that way. I tried to explain that in my original response but my response lacked clarity. Markov models are not about the percentage pf people with a scenario, it is about the direction of progression. Given the is a high proportion of people with live mets alone, or lung mets along, or peritoneal mets alone, these can't be seen as independent steps in a Markov chain. Our original data was perhaps poorly worded with the liver and lung mets group really mean to represent solitary mets inclusive of peritoneal. I did not list peritoneal because the vast majority of publications list liver, lung, bone or other, sometimes lymph, rarely peritoneal, sometimes brain. So for a markov model, it makes sense that, even if briefly, a single met occurs first, then multiple mets which usually includes bone plus one or more of the other solitary locations (which mets the Markov property) and then perhaps the less common mets like brain. So that is why we did the analysis the way we did. We cannot do the Markov analysis as requested because it does not meet the Markov properties but we believe tidying up our original language in the analysis provides the clarity the reviewer was looking for. Also, there is a lack of data for most other mets and data that is there is unreliable for peritoneal, lymph and lung mets (as per Uson et al 2020) hence the pooled data for multiple mets. Furthermore, there are survival data but no hazard ratios reported for much of the data which limits analysis. This approach is supported by Zhang 2023 (table 3) that outlines most probable metastatic sites for single or combinations for multiple. So solitary is predominantly liver and lung as per my original headings, multiple almost always involve bone with lung and liver, and rarely, brain joins multiple. SO this is what our Markov chain reflects.

Table 3

Patterns of distant metastases for the 2709 metastatic PDAC patients.

Sites of distant metastases	Overall (N=2709)
One site of distant metastasis	
Liver	1957 (72.2%)
Lung	251 (9.3%)
Bone	50 (1.8%)
Brain	2 (0.1%)
Two sites of distant metastasis	
Liver+lung	306 (11.3%)
Liver+bone	73 (2.7%)
Lung+bone	17 (0.6%)
Brain+bone	1 (0.0%)
Lung+brain	1 (0.0%)
Liver+brain	0 (0.0%)
Three sites of distant metastasis	
Liver+lung+bone	47 (1.7%)
Liver+lung+brain	2 (0.1%)
Liver+brain+bone	1 (0.0%)
Lung+brain+bone	0 (0.0%)
Four sites of distant metastasis	
Liver+lung+brain+bone	1 (0.0%)

Changes in text: We have refined the language to better reflect the multiple mets versus solitary mets as undertaken originally and have clarified the reasoning. We have simplified to single, multiple and multiple with brain to reflect the predominant data in the literature and hope this meets your needs for this query.