Reviewer(s)' Comments to Author:  
Referee: 1  
  
Comments to the Author  
In the manuscript by Bergman et al. the authors present a new computational model of tumor formation which includes also the tumor microenvironment, in particular the tumor infiltrating immune cells, immune evasion and effects of Tgfbeta signaling. Moreover, the model also takes into account that tumor cells can undergo EMT and display a mesenchymal-phenotype associated growth arrest. The authors find that with calculated rates of mutations, proliferation and apoptosis, EMT is regulating tumorigenesis and that the balance between EMT and immune-driven processes is orchestrating cancer-free survival. They support the model by using TCGA data of pancreatic cancer patients which confirms the synergistic effects. For the first time, a mathematical model is used that includes all these various aspects. This model is of great value for scientists of several research areas. Therefore the manuscript merits publication in the Journal of the Royal Society Interface. I summarized a few points that need to be addressed prior to publication:  
  
**1. In this mathematical model all events are happening at the very same time. This is maybe true for rates of mutation, proliferation, apoptosis and interaction with cells of the immune system. However, the concept of EMT is believed to be initiated as a rather late event during tumorigenesis, when a primary tumor has already formed during the transition from a benign to a malignant state. This idea is supported by the findings that in precancerous lesions, like in adenomas of the intestine or pancreatic intraepithelial neoplasia, mesenchymal cells a rarely found, whereas they are frequently identified in aggressive carcinomas. How does this fit to the proposed model? This temporal discrapency of current knowledge and the model should be discussed in more detail.**

We have clarified the focus of the model to be on tumor progression. We now begin with a benign growth that starts to acquire aggressive mutations and begins to undergo EMT as it works to escape the immune system.

**2. In line with this, the model is based on the time before a tumor/cancer is forming. However, to validate the model datasets from patients are used that monitor the period from the timepoint of cancer diagnosis onwards, represented as patients’ survival (which type is shown in Fig. 6: overall survival, relapse-free survival, metastasis-free survival?). From the perspective of cancer biology the process of tumor initiation is completely different to the one of tumor relapse. In the latter, dormant cancer cells which harbor already almost all mutations and MGA/MIE states are triggered to awake from dormancy.**

We will focus our data analysis on the time between initiation of treatment and relapse of the tumor.

**3. Tgfbeta signaling is known to have a dual role in tumorigenesis. It acts as a tumor suppressor and inducer of differentiation in early stages and as a tumor promoter in later stages while inducing EMT of epithelial cancer cells. This was analyzed in great detail by the Massagué group but is still not fully understood. In many tumors Tgfbeta signaling needs to be switched off by either mutating TGFBR2 or SMAD4 to escape this tumor suppressor effect. In lines 443-451 the authors discuss that low levels of Tgfbeta mediated effects act tumor suppressive in their model. However, the reasoning that this is due to plethoric functions of Tgfbeta and to off-target effects seems insufficient and inaccurate.**

In line with our earlier clarification, our tumor has already been initiated and is now beginning to progress. Thus, we can assume that the early-stage effects of TGFbeta have already been escaped from by the tumor and so TGFbeta takes on just two key roles in our model: recruiting/activating Tregs and inducing EMT. I am unsure what “off-target effects” refers to.

**4. In lines 343-345 the authors claim that MGA –mediated increase in Time to Cancer is acting on mutated and normal cells. This is very confusing since normal epithelial cells should not undergo EMT to face MGA and infiltrating normal stromal cells (CAFs, etc.) are already mesenchymal and proliferate unaffected by MGA. In other words, EMT and MGA is not happening in normal cells.**

Again per our clarification, all tissue cells are now assumed part of the tumor and thus can undergo aberrant EMT and experience the effects of MGA. There are no stromal cells modeled here.  
  
Referee: 2  
  
Comments to the Author  
The paper deals with the important topic of understanding the effect of immune system on EMT. The model appears new, and there is some comparision to real data. I have a number of comments, most of them may be due to unclear descriptions in the text. If these can be addressed and are not due to scientific flaws, but simply due to unclear statments, then the paper may be recommended for publication.  
  
  
- **better motivation for model assumptions. Especially, it shoudl be clearly stated for which assumptions there is a clear biological rational. Example: "Third, if the cell is mesenchymal, then the weight for proliferation is proportionally decreased" ... Give a reference, or at least clearly state this is an assumption but the real biology is not really known.**

Assumptions have been more clearly stated with a new table to list key assumptions. More references were given where needed.  
  
**'particularly in tumors originating in gastrointestinal and pancreatic tissues [7,8]. Do these references claim that particularly those cancers, rather than other cancers, aredriven by inflammation? I diud not get this impression.**

Clarified that there is at least evidence in these tumor types while the possibility still exists for others.  
  
**eq (2.5): for mathematical consistency, I assume sigma is small so that tau\_i is positive? Define tau in a way to cut off negative numbers.**

Reworded how tau\_i is described to better reflect what is going on in the model. Tau\_i can be negative and would lead towards a cell more likely undergoing MET. Since tau\_i can be negative, it is no longer defined as the TGFbeta received by cell i, but just a value to determine EMT/MET outcome.  
  
**sect 2.5 on parameter sensitivity: it is understandable that many parameters are not available, but a few more details on each assumption for each paramteter would be much more valuable than very generic statements (e.g. 'Where possible, these were informed by literature, elsewhere, we made informed approximations and chose relatively uninformative priors.'). Also why:  
'the variance for each was set to be twice the mean.' What is the motivation? And do you mean variance of standard deviation?**

Since many parameter values can only be assumed and the others found in the literature do not come with uncertainty, we simply choose a large spread for the possible parameter values relative to the mean value. Yes, standard deviation was meant.

**Fig 2. Could the authors be a bit more specific into how mu was calculated? It seems it is the average of sensitivities over 30 different points, each obtained from sampling from a normal distribution? So a zero sensitivity does not really mean no sensitivity, but one could have a linear response, so equally many and equally sized responses up and down, that average to zero? That would not be the standard sensitivity so I probably misunderstand the way mu is calculated**

It is the average of the *absolute value* of the changes. The wording has been updated to better emphasize this.

**Section 3.5: in principle an interesting section comparing to data. But to me it is totally irreducible and unclear what the authors exactly did here. It is nice that the code is there but the description in the text should also allow one to reproduce exactly what the authors did. How was this data compared to the model? For this and the model code it would be nice to have a clear description of the code, maybe in the SI.**  
  
The data was only compared in a qualitative way showing that EMT and inflammation taken together better predict patient outcome than either alone.

Referee: 3  
  
Comments to the Author  
Review of:   
Modeling the competing effects of the immune system and EMT on tumor development  
  
By Bergman et al  
  
The authors present a model that involves tissue and immune interactions during carcinogenesis with a focus on EMT. The model is non-spatial agent-based with continuous equations for the immune components. They model inflammation in the system by changing the parameter values for some of the immune components. They find that increasing mesenchymal evasion, decreasing mesenchymal growth arrest, and increasing TGF-beta production all lead to a shorter time to cancer. However, low periods of inflammation can lead to an optimal value for growth arrest. They investigate signatures for EMT and inflammation in pancreatic cancer data from TCGA and find that there is decreased survival when both inflammation and EMT signatures are present over EMT alone.  
  
Whilst EMT associated with cell invasiveness and metastasis is well studied, the idea of EMT at the stages of cancer initiation is a new perspective to me. The authors cite reference [10], which indicates that these are partial transitions occurring in cells that might not be associated with a typical concept of a mesenchymal phenotype, but more of loss of tissue structure probable during the early cancer initiation stage. However, with only this one reference and lack of clarification on the ideas, it is hard to assess the biological significance. In addition, the model then ignores this nuance in EMT status and assigns either an epithelial or mesenchymal phenotype to the cell. This causes more confusion and disrupts the earlier assumption of plasticity and partial transition that was needed to make sense of early EMT. Furthermore, this early EMT and the more familiar concept of EMT at later stages is not necessarily related. And if they are on the same spectrum of traits, it could be argued that they represent very different processes at early and late time points, especially when it comes to recognition by and response of the immune system and the importance of spatial structure at later time points.   
  
That said, getting a better handle on inflammation, immune responses, and early tissue changes prior to cancer is an intriguing and worthwhile pursuit, and relating these components to each other during later stages is also needed. Given this setup, there appears to be some interesting non-linear, non-intuitive behavior in the parameter space, which could have applications in better understanding the complex behavior between tumors, their environment, and the immune system. However, in the end, this research rests on biologically unclear assumptions, which make significance of the results questionable.   
  
  
MAJOR ISSUES:  
**1. EMT is not generally considered to be important at early stages of cancer progression, especially in the pre-cancerous state. If there is evidence to the contrary, it would be beneficial to discuss this in the introduction to set up the problem for the reader. More to this point, the authors neglect any role of space in the assumptions, which is fine for the pre-cancerous state. However, if EMT is being induced by hypoxia, and occurring at later time points as a reaction to the lack of resources. It would be beneficial to address this assumption of the model, and to better connect the early stage cancer initiation concepts to the later stage TCGA data analysis linked to survival.**

As per our response to Referee 1, we have clarified that our model is tracking tumor progression and so we are not focusing on the pre-cancerous state. We have also added some discussion of our lack of spatial modeling and the resulting assumptions we must make.

**2. Fig. 1A is pretty confusing. It seems that the vertical axis represents different mutations in the pathways listed, but it is not clear how the lack of mutation transitions to the epithelial or mesenchymal cancerous state. It is not at all clear what the boxes represent horizontally or why this confusing figure is repeated in the insert. It is also too dark and small to read the text in the insert. Also, in the text, it states that red is mutated, while in the caption, it states blue/red is with/without mutations. Cleaning up this figure would certainly help orient the reader.**

Fixed the typo in the caption. Also, added more detail to the caption to better explain the figure.

**3. Probably some of the confusion in setting up this work is the lack of clarity in the terminology used. EMT, plasticity, stemness are all ambiguous terms that might need more explicit definitions, especially when looking at cancer initiation due to the fact that observation at this stage is also challenging to observe and quantify. Perhaps more precise language and clarification will help.**

Clarified that our model does not address initiation, but rather just progression. Thus, these terms are no longer being stretched to cover different periods of tumor growth.

**4. The authors use a certain threshold of mutations to determine time to cancer. There is, however, evidence that neutral evolution can cause a large mutational burden without any incidence of cancer. See work by Martincorena in Science 2015 and 2018.**

Our model focuses on pathway mutations to three key pathways known to be hallmarks of cancer. Any mutations to these pathways will not only increase the mutational burden but increase the likelihood of cancer. We ignore any other mutations.

MINOR ISSUES:  
  
**1. Line 87: What do you mean by synergistic effects predicted by the model? Synergism between what attributes?**

Synergism between EMT and inflammation.

**2. Line 100: What do you mean by score? Is this a mutational status? And is it the total mutational sum or of a particular kind?**

EMT score. Clarified in the text.

**3. Incude reference that states tgfb is produced by Tregs to induce mesenchymal states.**

Shi C, Chen Y, Chen Y, Yang Y, Bing W, Qi J. CD4+ CD25+ regulatory T cells promote hepatocellular carcinoma invasion via TGF-β1-induced epithelial–mesenchymal transition. OncoTargets and therapy. 2019;12:279.

**4. What is the “warm-up” period for? This is not clear in the text.**

The warmup is to ensure that a steady-state has been reached given the model parameters.

**5. It is not clear how the inflammation cycling occurs.**

It happens in a fixed cycling scheme where it is high for some number of days and low for some other number of days. These can vary between patients.

**6. Eqn 2.4 is not well explained as far as the variables. It seems that there should be a minus sign somewhere if the equation is describing the weight of quiescence. Also, here the authors use the term quiescence even though they state that they will avoid it due to controversy earlier (by the way, I don’t find it controversial, but it helps to be consistent).**

The proliferation that mesenchymal cells lost is moved to their rest weight. Hence, the addition of this quantity to the base rest weight of 1. All references to quiescence have been removed and we now only use the term rest (or G0).

**7. Eqn 2.5: Why use a Hill function to describe the amount of TGFb? It makes sense to consider the effect of TGFb as a saturating function, but would you explain why to use this form for the quantity available for each cell?**

Each cell can only absorb so much TGFbeta in a given time interval.

**8. Lines 209-210: what is a “tumorigenic event” in the model?**

This is when 50% of the tissue cells have undergone a pathway mutation. The wording has been changed to clarify this connection.

**9. Line 235 & Line 237: what is msigdb and S?**

Fuller explanations given.

**10. Please include units for the values in Table 2. It would also help to include references for each of the parameter values.**

Units included.