BMJ Open

Diminishing returns along the road to translation: a systematic review T-VEC's preclinical to clinical development trajectory

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-029475
Article Type:	Research
Date Submitted by the Author:	30-Jan-2019
Complete List of Authors:	Lalu, Manoj; Ottawa Hospital Research Institute, Clinical Epidemiology Leung, Garvin; Ottawa Hospital Research Institute, Clinical Epidemiology Dong, Yuan Yi; Ottawa Hospital Research Institute, Clinical Epidemiology Montroy, Joshua; Ottawa Hospital Research Institute, Clinical Epidemiology Butler, Claire; Ottawa Hospital Research Institute, Clinical Epidemiology Auer, Rebecca; Ottawa Hospital Research Institute, Clinical Epidemiology Fergusson, Dean; Ottawa Hospital Research Institute, Medicine
Keywords:	TVEC, oncolytic virus, cancer, translation, review

Diminishing returns along the road to translation: a systematic review T-VEC's preclinical to clinical development trajectory

Manoj M. Lalu† MD, PhD, FRCPC[1,2,3,4], Garvin J. Leung† MPH[2,5], Yuan Yi Dong[2,5], Joshua Montroy MSc,[2] Claire Butler[2], Rebecca C. Auer MD, MSc, FRCSC[7,8], Dean A. Fergusson PhD, MHA*[2,6,7]

[1]Department of Anesthesiology and Pain Medicine, The Ottawa Hospital, University of Ottawa, Ottawa, ON, Canada
[2]Blueprint Translational Research Group, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
[3] Regenerative Medicine Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
[4] Department of Cellular and Molecular Medicine, University of Ottawa , Ottawa, ON, Canada
[5]Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada
[6]School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada
[7]Department of Surgery, The Ottawa Hospital, University of Ottawa, Ottawa, ON, Canada
[8]Cancer Therapeutics Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
† These authors contributed equally to this work

Emails: MML: mlalu@toh.ca; GJL: garvin.leung@uottawa.ca; YYD: ydong044@uottawa.ca; JM: jmontroy@ohri.ca; CB: clbutler@toh.ca; RCA: rauer@toh.ca; DAF: dafergusson@ohri.ca

*Corresponding author:
Dean A. Fergusson email: dafergusson@ohri.ca
Centre for Practice-Changing Research,
Office L1298a
501 Smyth Road, Box 201B
Ottawa, Ontario, Canada K1H 8L6
Tel. 1-613-737-8480
Fax. 1-613-739-6938

Running Head: The efficacy of T-VEC: A systematic review

Word count: 2631

ABSTRACT

Objective: This study aimed to conducted a systematic review of preclinical and clinical evidence to map the successful trajectory of Talimogene laherparepvec (T-VEC), from the bench to the clinic.

Design: This study was a systematic review. The primary outcome of interest was the efficacy of treatment, determined by complete response. Abstract and full-text selection as well as data extraction was done by two independent reviewers. The Cochrane risk of bias tool was used to assess the risk of bias in studies.

Setting: Embase, Embase Classic, and OvidMedline were searched from inception until May 2016 to assess its development trajectory to approval in 2015.

Participants: Preclinical and clinical controlled comparison studies, as well as observational studies.

Interventions: T-VEC for treatment of any malignancy.

Results: 8,852 records were screened and five preclinical (n=150 animals) and seven clinical studies (n=589 patients) were included. We saw large decreases in T-VEC's efficacy as studies moved from the laboratory to patients, and as studies became more methodologically rigorous. Preclinical studies reported complete regression rates up to 100% for injected tumors and 80% for contralateral tumors, while the highest degree of efficacy seen in the clinical setting was a 24% complete response rate, with one study experiencing a complete response rate of 0%. We were unable to reliably assess safety due to the lack of reporting, as well as the heterogeneity seen in adverse event definitions. All preclinical studies had high or unclear risk of bias, and all clinical studies were at a high risk of bias in at least one domain.

Conclusions: Our findings illustrate that even successful biotherapeutics may not demonstrate a clear translational road map. This emphasizes the need to consider increasing rigour along the translational pathway.

PROSPERO Registration Number: CRD42016043541

Keywords: TVEC, oncolytic virus, cancer, translation, review

Strengths and limitations of this study

- Comprehensive, up-to-date review of the efficacy and safety of TVEC
- Threats to both internal validity and construct validity were performed
- Reporting of methods and findings was incomplete in most of the studies included
- Poor reporting and study design are major contributors to the ongoing reproducibility crisis in preclinical research

BACKGROUND

Preclinical research receives approximately half of the world's biomedical research funding, yet very few of its findings translate clinically. This represents an enormous waste of resources with an estimated 28 billion dollars per year in the US alone being spent on biomedical research which is not reproducible and therefore not translatable.(1) One study found that only 5% of highly efficacious preclinical therapeutics were clinically translated.(2) These successes often take almost twenty years to become successfully translated. (2, 3) Given the high failure rate in translating therapies and significant time-lags, it is crucial we evaluate and learn from the few agents that have successfully crossed the preclinical-to-clinical bridge in order to learn from and replicate their success.

Thus, we conducted a comprehensive evaluation of available evidence supporting the successful translation of Talimogene laherparepvec (T-VEC). T-VEC is a modified HSV-1 virus produced by *Amgen* and it is the first, and only, FDA approved oncolytic virus therapy; it is currently approved to treat advanced melanoma.(4) Oncolytic viruses are an emerging cancer therapy that work by preferentially targeting and infecting cancer cells.(4) Upon infection, oncolytic viruses can induce an anti-tumor immune response that reduces tumor burden.

Through a careful evaluation of T-VEC development we hoped to identify factors that may contribute to bench-to-bedside success. This may serve an exemplar for other therapies as they move along the translational continuum. Thus, the purpose of this systematic review was to map the successful preclinical to clinical trajectory of T-VEC to inform the development paths of new biotherapeutics.

METHODS

Our review was registered in full on PROSPERO, the international prospective register of systematic reviews (no. CRD42016043541). The review is reported in accordance to the PRISMA guidelines.(5)

Eligibility Criteria

We included all clinical and preclinical *in vivo* controlled comparison studies of TVEC for treatment of any malignancy (randomized, pseudo-randomized, and non-randomized studies), as well as observational studies such as case-control, case-series and case reports. Studies reporting only *ex vivo* or *in vitro* experiments were excluded. For both preclinical and clinical studies, we included studies that administered TVEC as a monotherapy or in combination with other therapies for treatment of malignancy. We had no exclusions on comparison treatments, which include standard line therapy or no treatment.

Outcomes

The primary outcome of interest was the efficacy of treatment. Our primary indicator of efficacy was complete response. Other measures of efficacy such as survival, response rates (durable, partial, objective), time to treatment failure, and disease stability were also collected. Such measures were based on the Response Evaluation Criteria in Solid Tumors (RECIST) Guidelines.(6) In preclinical studies, additional measures of efficacy such as changes in mean tumor volume and number of lesions were collected. The secondary outcome of interest was safety, for which we collected data on all adverse events in preclinical and clinical studies.

Literature Search

In collaboration with a medical information specialist (Risa Shorr, Learning Services, The Ottawa Hospital) a search strategy was designed to identify all relevant preclinical and clinical studies. Searches were conducted in the following databases: Embase, Embase Classic, and OvidMedline from inception until May 2016. This time frame was chosen to ensure all published studies that contributed T-VECs FDA approval in 2015 were included. Search terms included: Talimogen laherparepvec, Tvec, OncoVEX and Imlygic. Additional terms pertaining to preclinical studies (e.g. animal experiment/model) and oncology (e.g. cancer, neoplasm, oncolytic virus) were also included. Studies were also screened for inclusion based on reference tracking, by scanning the bibliography of included primary studies and relevant review articles. We did not impose any restrictions on language or publication type. The finalized search strategy can be found in Appendix 1.

Study Selection Process

Studies identified by our literature search were collated and duplicates were removed. Titles and abstracts were independently screened for inclusion by two reviewers using DistillerSR (Evidence Partners, Ottawa, ON). Those deemed potentially relevant were recorded, and full-text articles were obtained. The same reviewers screened full articles for final eligibility. Disagreements at any stage were resolved by discussion or by consultation with a senior team member when necessary. The study selection process was documented using a PRISMA flow diagram (Figure 1).

Data Extraction

All data extraction was completed independently and in duplicate, using a standardized and piloted data extraction form, with disagreements resolved as mentioned above. Data pertaining to general

and intervention characteristics of the included studies were extracted (e.g. study design, country, type of malignancy, dosing of intervention and comparator treatments). For clinical studies, data was collected on patient characteristics (e.g. age, sex, cancer staging, HSV status). For preclinical studies, characteristics on the animal model were extracted (e.g. type of species, cell line used, disease induction method, age, sex, weight).

Risk of bias – assessment to risk of internal validity

Clinical studies that met inclusion criteria were assessed for risk of bias in duplicate, according to the recommended methodology of the Cochrane Collaboration. Five types of biases (selection, performance, detection, attrition, reporting biases) were assessed using six domains: randomization, allocation concealment, blinding of participants/personnel, outcome assessment blinding, incomplete outcome reporting, and selective outcome reporting. Additional domains assessed for risk of bias were: i) reported conflicts of interest, ii.) sample size calculation, and iii.) funding. Each domain was given a score of "high", "unclear", or "low" risk of bias for each study. Risk of bias assessment for preclinical studies were assessed using a modified Cochrane Risk of Bias tool and assessed the same domains as indicated for clinical studies.

Assessment of threats to construct validity

Construct validity is the concept in how much a preclinical experiment (i.e. animal studies) corresponds to the clinical entity it is intended to model. There are various threats to construct validity that can be introduced from the preclinical study design. The items evaluated in duplicate for each preclinical study include: i.) use of adult animals, ii.) use of animals with advanced stage disease (defined as the presence of multiple visceral lesions and/or clinical/histological signs of malignant progression), iii.) immune status of animals to HSV, iv.) whether a xenograft model was

used, and v.) the use of a humanized immune system model. Each of these items was given a score of "yes", "no", or "unclear" for every preclinical study.

Statistical Analysis

Efficacy was expressed as simple proportions. To assess the continuity between preclinical and clinical studies, the efficacy of studies was plotted as percentage response.

Deviations from Protocol

We were unable to assess safety as we could not acquire patient-level safety data. Furthermore, our primary efficacy outcome stated in protocol was durable response rate. However, this was changed to complete response as most clinical studies did not report durable response. Subgroup analyses, meta-analyses, Egger's test, and pooling of data could not be conducted due to the limited available data.

Patient and Public Involvement

Patients and the public were not involved in this research.

RESULTS

Upon removal of duplicates, a total of 8,852 references were identified by the electronic search. During the review of titles and abstracts, 7,890 references were excluded. Following full text screening, a total of seven clinical studies,(7-13) and five preclinical studies(14-18) were included in our review (Figure 1).

Characteristics of Included Trials

Characteristics of included studies are shown in Table 1. Preclinical studies were published between 2003 and 2016 and sample sizes ranged from 20 to 90. Of the five preclinical studies,

three used a lymphoma model, one used a colorectal model, and one used a melanoma model. All studies were performed in mice. The duration of follow-up was reported by two studies and ranged from 10 days to 35 days. The dose of TVEC used ranged from $3x10^4$ plaque forming units (PFU) to $5x10^6$ PFU. One frequency of TVEC administration varied from, every three days for one week, every three days for nine days, a single dose given only once, and every other day for five days. Specific details of study and intervention characteristics for each preclinical study can be found in Appendix 2.

Clinical studies were published between 2006 and 2016 and took place in seven countries. Sample sizes ranged from 17 to 295. Of the seven clinical studies, four were in melanoma patients, one was in pancreatic cancer patients, one in head and neck cancer patients and one studied breast, colorectal, melanoma and head and neck cancer patients. Six were either Phase I or II, and one trial was a Phase III evaluation. The primary outcome was efficacy in two studies, safety in three studies and a combination of efficacy and safety in the other two studies. The duration of follow-up ranged from six weeks to 44 months.

TVEC was administered alone in four studies, while it was administered adjuvant to chemotherapy in 3 studies. The dose of TVEC administered ranged from 10^4 PFU/mL to 10^8 PFU/mL. In the large, Phase III study, TVEC was administered at ≤4mL x 10^6 PFU/mL once, and then three weeks later, ≤4mL 10^8 PFU/mL was administered every two weeks for a median of 23 weeks. A similar dosing regimen was used in three other trials. The other three trials were dose-finding in nature and had multiple trial arms receiving increasing doses of TVEC. In-depth study details, as well as participant and intervention details for each study can be found in Appendix 2.

Efficacy of Treatment

Treatment efficacy for each study is summarized in Table 1 and Figure 2. Preclinical studies reported complete regression rates up to 100% for injected tumors and 80% for contralateral tumors. In comparison, the first published Phase I T-VEC clinical trial reported a complete response of 0% for cutaneous lesions caused by malignancies of head and neck, breast, colorectal, and melanoma [4]. Of the multiple malignancies treated, melanoma had the best response in this trial. Subsequent Phase I/II melanoma trials were then conducted and demonstrated complete response rates of 20-22%. This was followed by the Phase III OPTIM melanoma trial, which had a complete response rate of 10.8%.[5–7] Studies involving non-melanoma cancers varied with efficacies between 0-24%.[8,9]

Safety of treatment

We attempted to assess safety, however we were unable to obtain patient level data from any of the studies. The definitions of adverse events, and the manner in which they were classified, was found to be highly heterogenous across studies, therefore we were unable to pool adverse events or interpret findings reliably.

Validity Assessments

Construct validity, the concept of how well an animal model represents the clinical entity it is intended to mimic, was first assessed through the following domains: the use of appropriately-aged mice, advanced stage of disease, HSV-immunity, and types of mouse models. None of the preclinical studies fully reported or used methodologies to reduce threats to construct validity domains (Table 2). No studies declared using adult animal models, no studies used animals with late stage disease, only one study used animals immune to HSV, no studies used a xenograft model, and no studies reported using an animal model with a humanized immune system.

We also assessed internal validity (i.e. risk of bias) and found that all preclinical studies had high or unclear risk of bias across the assessed domains: randomization sequence, allocation concealment, blinding, incomplete reporting, sample size calculation, and funding source (Table 3). For clinical studies, early phase trials had high or unclear risk of bias across at least six of nine domains whereas the more robust Phase III OPTIM trial had the lowest risk of bias and also the lowest efficacy of any of the published melanoma clinical trials (Table 4). Reporting of key methodological elements was lacking.

DISCUSSION

We hoped to synthesize a clear road map of T-VEC's translation in the published literature to follow the journey a successful biotherapeutic travels. Yet, we were unable to paint a clear picture of how the evidence was utilized in proceeding to melanoma clinical trials. Rather, our assessment uncovered a clear disconnect between *in vivo* preclinical and clinical findings. Furthermore, the road map was plagued with poor reporting, high risk of bias, and insufficient data along the translational path. Overall, we were surprised by the pace and magnitude of diminishing efficacy as T-VEC moved from bench to bedside and then towards later phase clinical trials (i.e. Phase I to III).

While many novel therapeutics are under intellectual property rights, details of study design and results should be transparently reported for scientists, clinicians, and patients to evaluate findings. The fact that the only FDA approved oncolytic virus therapy is not clearly reported illustrates the issues plaguing the success of cancer therapeutics. Nonetheless, T-VEC has shown some efficacy in treating refractory melanoma and numerous clinical trials are underway to assess its use in combination with other cancer regimens and in treating other malignancies. While we recognize that translation is not a linear process, we should observe consistent and coherent patterns. Moving

forward, we suggest that preclinical and clinical studies for emerging therapies should be fully reported and attention should be given to validities in order to develop more precise estimates of effect early in development. We believe these steps will provide unbiased and valuable information that will ultimately provide patients with potentially more efficacious cancer therapies and protect them against needless evaluation.

Perhaps the largest discrepancy noted was that only a single preclinical study used a melanoma model, whereas 5/7 clinical studies administered T-VEC to melanoma patients. Conversely, lymphoma, which was used in three preclinical studies, was not assessed in clinical studies. Interestingly, our subsequent searches found that Amgen's FDA filing (STN# 125518.000) for T-VEC did not appear to report on any *in vivo* melanoma models, whereas the EMA report did (EMA/734400/2015). Thus, the majority of animal models were off-target from the malignancies studied in clinical trials and may have poorly represented melanoma in the clinical setting. Coupled with these findings was the fact that the majority of our studies were found to be at a high risk of bias.

Such threats to internal validity can bias results and may help explain T-VEC's superior preclinical efficacy compared to later phase clinical trials. A lack of randomization and blinding in preclinical studies has been associated with inflated effect sizes,(19, 20) thus this may partially explain the preclinical to clinical discrepancy of T-VEC.

Reporting of methods and findings was incomplete in most of the studies included. Only one full preclinical article on T-VEC was published, and solely aggregate patient data for later phase trials was available. Poor reporting and study design are major contributors to the ongoing reproducibility crisis in preclinical research.(21) Thus, in hopes of presenting a clearer picture of T-VEC's successful translation, we contacted *Amgen* to obtain preclinical *in vivo* melanoma data,

patient-level safety data, and any additional efficacy data. Patient-level data would afford the ability to combine data across T-VEC's clinical development and also provide clarification into the categorization of adverse events. Recently, release of individual patient data to third parties has been advocated by the Institute of Medicine, journal editors, and others as it enhances transparency, enables re-analyses of data, and helps address reproducibility.(22) However, *Amgen* was unwilling to enter a data sharing agreement, as they stated that there was little value to compel a transparent data release for our proposed analyses. This lack of transparency and incomplete reporting is disappointing, especially considering that it was *Amgen* that previously fingered poor reporting as contributing to its own failure to reproduce 47 of 53 high-impact preclinical cancer studies.(23) Their findings fuelled a call by the NIH and other stakeholders to enhance the reproducibility and transparency of preclinical research.(24)

CONCLUSIONS

The findings from our systematic review demonstrate that even successful biotherapeutics may not demonstrate a clear translational road map. The magnitude of efficacy of T-VEC demonstrated in preclinical studies was considerably larger when T-VEC was moved to the clinic, and the most methodologically rigorous trial included in our review demonstrated the smallest degree of efficacy. Methodologically rigorous studies should be performed earlier on in the translational pathway, which may help to get a realistic estimate of treatment efficacy prior to clinical translation.

DECLARATIONS

Ethics approval: Not applicable.

Consent for publication: Not applicable.

Availability of data: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding: Biotherapeutics for Cancer Treatment (BioCanRx) supported the conduct of this study by a Catalyst grant. BioCanRx is a Government of Canada funded Networks of Centres of Excellence and was not involved in any other aspect of the project, such as the design of the project's protocol and analysis plan, the collection of data and analyses. CB was also supported by a BioCanRx studentship. ML is supported by The Ottawa Hospital Anesthesia Alternate Funds Association and the Scholarship Protected Time Program, Department of Anesthesiology and Pain Medicine, uOttawa.

Author contributions: MML and DAF conceptualized the study. MML, RA, DAF, and GJL contributed to the study design. GJL, YYD, and CB conducted data extraction. All authors analysed and interpreted the data. MML and GJL were responsible for drafting the manuscript. All authors critically reviewed the manuscript and provided intellectual content. All authors approve the final version of the manuscript.

Acknowledgments: We thank Risa Shorr (Ottawa Hospital Research Institute librarian) for systematic search assistance. MML is supported by The Ottawa Hospital Anesthesia Alternate Funds Association and the Scholarship Protected Time Program, Department of Anesthesiology and Pain Medicine, uOttawa.

REFERENCES

1. Freedman LP, Cockburn IM, Simcoe TS. The Economics of Reproducibility in Preclinical Research. PLoS biology. 2015;13(6):e1002165.
2. Contopoulos-Ioannidis DG, Ntzani E, Ioannidis JP. Translation of highly promising basic science research into clinical applications. The American journal of medicine. 2003;114(6):477-84.
3. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. Journal of the Royal Society of Medicine. 2011;104(12):510-20.
4. Rehman H, Silk AW, Kane MP, Kaufman HL. Into the clinic: Talimogene laherparepvec (T-VEC), a first-in-class intratumoral oncolytic viral therapy. J Immunother Cancer. 2016;4:53.
5. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009;62(10):1006-12.
6. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47.
7. Hu JC, Coffin RS, Davis CJ, Graham NJ, Groves N, Guest PJ, et al. A phase I study of OncoVEXGM-CSF, a second-generation oncolytic herpes simplex virus expressing granulocyte macrophage colony-stimulating factor. Clinical cancer research : an official journal of the American Association for Cancer Research. 2006;12(22):6737-47.
8. Senzer NN, Kaufman HL, Amatruda T, Nemunaitis M, Reid T, Daniels G, et al. Phase II clinical trial of a granulocyte-macrophage colony-stimulating factor-encoding, second-generation oncolytic herpesvirus in patients with unresectable metastatic melanoma. J Clin Oncol. 2009;27(34):5763-71.
9. Andtbacka RHI, Kaufman HL, Collichio F, Amatruda T, Senzer N, Chesney J, et al. Talimogene Laherparepvec Improves Durable Response Rate in Patients With Advanced Melanoma. Journal of Clinical Oncology. 2015;33(25):2780-8.
10. Long GV, Dummer R, Ribas A, Puzanov I, Michielin O, Vanderwalde A, et al. 24LBA Safety data from the phase 1b part of the MASTERKEY-265 study combining talimogene laherparepvec (T-VEC) and pembrolizumab for unresectable stage IIIB-IV melanoma. European Journal of Cancer. 2015;51:S722.
11. Puzanov I, Milhem MM, Minor D, Hamid O, Li A, Chen L, et al. Talimogene Laherparepvec in Combination With Ipilimumab in Previously Untreated, Unresectable Stage IIIB-IV Melanoma. J Clin Oncol. 2016;34(22):2619-26.
12. Harrington KJ, Hingorani M, Tanay MA, Hickey J, Bhide SA, Clarke PM, et al. Phase I/II study of oncolytic HSV GM-CSF in combination with radiotherapy and cisplatin in untreated stage III/IV squamous cell cancer of the head and neck. Clinical cancer research : an official journal of the American Association for Cancer Research. 2010;16(15):4005-15.
13. Chang KJ, Senzer NN, Binmoeller K, Goldsweig H, Coffin R. Phase I dose-escalation study of talimogene laherparepvec (T-VEC) for advanced pancreatic cancer (ca). Journal of Clinical Oncology. 2012;30(15_suppl):e14546-e.
14. Liu BL, Robinson M, Han ZQ, Branston RH, English C, Reay P, et al. ICP34.5 deleted herpes simplex virus with enhanced oncolytic, immune stimulating, and anti-tumour properties. Gene Ther. 2003;10(4):292-303.
15. Piasecki J, Tiep L, Zhou J, Beers C. Talilmogene laherparepvec generates systemic T-cell-mediated anti-tumor immunity. Journal for ImmunoTherapy of Cancer. 2013;1(1):P198.
16. Cooke K, Rottman J, Zhan J, Mitchell P, Ikotun O, Yerby B, et al. Oncovex MGM-CSF –mediated regression of contralateral (non-injected) tumors in the A20 murine lymphoma model does not involve direct viral oncolysis. Journal for Immunotherapy of Cancer. 2015;3(Suppl 2):P336-P.

17. Piasecki J, le T, Ponce R, Beers C. Abstract 258: Talilmogene laherparepvec increases the anti-tumor efficacy of the anti-PD-1 immune checkpoint blockade. Cancer Research. 2015;75(15 Supplement):258-.
18. Cooke K, Estrada J, Zhan J, Mitchell P, Bulliard Y, Beltran PJ. Abstract 2351: Development of a B16F10 cell line expressing mNectin1 to study the activity of OncoVEXmGM-CSF in murine syngeneic melanoma models. Cancer Research. 2016;76(14 Supplement):2351-.
19. Hirst JA, Howick J, Aronson JK, Roberts N, Perera R, Koshiaris C, et al. The need for randomization in animal trials: an overview of systematic reviews. PLoS One. 2014;9(6):e98856.
20. Macleod MR, van der Worp HB, Sena ES, Howells DW, Dirnagl U, Donnan GA. Evidence for the efficacy of NXY-059 in experimental focal cerebral ischaemia is confounded by study quality. Stroke. 2008;39(10):2824-9.
21. Landis SC, Amara SG, Asadullah K, Austin CP, Blumenstein R, Bradley EW, et al. A call for transparent reporting to optimize the predictive value of preclinical research. Nature. 2012;490(7419):187-91.
22. Zarin DA, Tse T. Sharing Individual Participant Data (IPD) within the Context of the Trial Reporting System (TRS). PLoS Med. 2016;13(1):e1001946.
23. Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. Nature. 2012;483(7391):531-3.
24. Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. Nature. 2014;505(7485):612-3.

Figure Legends

Figure 1. Study selection flow diagram

Figure 2. Preclinical and clinical efficacy of T-VEC. Four preclinical studies using mice demonstrated efficacy rates of 20-100% and clinical studies (3 melanoma and 2 mixed malignancy studies) demonstrated efficacy rates from 0-23.5%. Non-melanoma/mixed studies are represented by blue bars whereas melanoma studies are represented by orange bars. Where possible, complete regression rates of contralateral tumors for mice were used in preclinical studies and complete response was used for clinical studies. Efficacy rates decrease in the preclinical to clinical translation and upon more rigorous study design in later phase clinical trials. If possible, error bars are plotted and represent 95% confidence intervals. Some studies were not included in this analysis due to no reporting of outcome for CR.

Table 1A. *Study characteristics of included preclinical studies of T-VEC.*

Preclinical Study	Treatment	Total Number of Animals Used	Type of Cancer/Model	Efficacy Measures*	Risk of Bias (/9**)
Liu, 2003[14]	T-VEC; HSV1 wildtype immunization	90	Lymphoma (A20 murine lymphoma mouse model)	CR: 100% (n=10) (injected)	9
Piasecki, 2013[15]	T-VEC	NR	Lymphoma (A20 murine lymphoma mouse model)	CR: 70-100% of injected, 50-60% of contralateral	9
Piasecki, 2015[17]	T-VEC + Anti-PD-1	NR	Colorectal (MC-38 colon carcinoma mouse model)	CR: 80.0% (44.2-96.5%) (injected) n=10 CR: 20.0% (3.5-55.8%) (contralateral) n=10	9
Cooke, 2015[16]	T-VEC	40	Lymphoma (A20 murine lymphoma mouse model)	CR: 100% (65.5-100%) (injected) n=10 CR: 50% (23.7-76.3%) (contralateral)	9
Cooke, 2016[18]	T-VEC	20	Melanoma (B16F10 melanoma model)	NR – statistically significant tumor reduction and survival noted.	9

Table 2B. *Study characteristics of included clinical studies of T-VEC.*

Clinical Study	Treatment	Total N	Type of Cancer	Efficacy Measures*	Risk of Bias (/9**)
Hu, 2006[7] *Non-controlled Phase I*	T-VEC	30 (9 melanoma)	Breast, colorectal, melanoma, head and neck	CR: 0% (0-14.1%) PR: 0% (0-14.1%)	7
Senzer, 2009[8] *Non-controlled Phase II*	T-VEC	50	Melanoma	OR: 26.0% (15.1-40.6%) CR: 20.0% (10.5-34.1%) PR: 10.0% (3.7-22.6%)	7
Harrrington, 2010[12] *Phase I/II*	T-VEC + cisplatin	17	Head and neck	CR: 23.5% (7.8-50.2%) PR: 58.8% (33.5-80.6%) OR: 82.4% (55.8-95.3%)	6
Chang, 2012[13] *Phase I*	T-VEC	17	Pancreatic	OR: 0% (0-22.9%)	6
Andtbacka, 2015[9] *Phase III*	T-VEC	295	Melanoma	DR: 16.3% (12.1-20.5%) OR: 26.4% (21.4-31.5%) PR: 15.6% (11.7-20.3%) CR: 10.8% (7.6-15.1%)	3

	GM-CSF (control)	141		DR: 2.1% (0-4.5%) OR: 5.7% (1.9-9.5%) PR: 5.0% (1.3-8.5%) CR: < 1%	
Long, 2015[10] *Phase Ib*	T-VEC + pembrolizumab	21	Melanoma	-	6
Puzanov, 2016[11] *Phase Ib*	T-VEC + IPI	18	Melanoma	DR: 44.4% (22.4-68.7%) OR: 50.0% (29.0-70.9%) CR: 22.2% (7.4-48.1%) PR: 27.8% (10.7-53.6%)	6

* DR – durable response; OR – objective response; CR – complete response/complete regression; PR – partial response; DR/OR/CR/PR definitions were based on RECIST guidelines for clinical studies. **Total number of domains that were assessed a score of high risk or unclear (maximum = 9).

Table 3. Construct validity assessment for preclinical studies

Author, Year	Adult Used	Animals with Advanced Stage Disease	Animals Immune to HSV	Xenograft Model Used	Used Model with a Humanized Immune System
Cooke, 2016	Unclear	No	Unclear	No	Unclear
Cooke, 2015	Unclear	No	Unclear	No	Unclear
Piasecki, 2015	Unclear	Unclear	Unclear	No	Unclear
Piasecki, 2013	Unclear	Unclear	Unclear	No	Unclear
Liu, 2003	Unclear	No	Yes	No	Unclear

Table 4. Risk of bias assessment for preclinical studies

Author, Year	Random Sequence Generation	Allocation Concealment	Blinding of Personnel	Blinded Outcome Assessment	Incomplete Outcomes Addressed	Selective Outcome Reporting	Conflicts of Interest	A Priori Sample Size Calculation	Funding
Cooke, 2016	High Risk	High Risk	Unclear	Unclear	High Risk	High Risk	Unclear	Unclear	High Risk
Cooke, 2015	High Risk	High Risk	Unclear	Unclear	High Risk	High Risk	Unclear	Unclear	High Risk
Piasecki, 2015	Unclear	Unclear	Unclear	Unclear	High Risk	High Risk	Unclear	Unclear	High Risk
Piasecki, 2013	Unclear	Unclear	Unclear	Unclear	High Risk	High Risk	Unclear	Unclear	High Risk

Liu, 2003	Unclear	Unclear	Unclear	Unclear	High Risk	Unclear	Unclear	Unclear	High Risk

Table 5. Risk of bias assessment for clinical studies

Author, Year	Random Sequence Generation	Allocation Concealment	Blinding of Participants and Personnel	Blinding of Outcome Assessors	Incomplete Outcome Data Addressed	Selective Reporting	Conflicts of Interest	Funding	Sample Size Calculation
Andtbacka, 2015	Low Risk	Low Risk	High Risk	Low Risk	Low Risk	Low Risk	High Risk	High Risk	Low Risk
Long, 2015	High Risk	N/A	High Risk	Unclear	Low Risk	Low Risk	High Risk	High Risk	Unclear
Puzanov, 2016	High Risk	N/A	High Risk	Unclear	Low Risk	Low Risk	High Risk	High Risk	Unclear
Chang, 2012	High Risk	N/A	High Risk	Unclear	Low Risk	Low Risk	High Risk	High Risk	Unclear
Harrington, 2010	High Risk	N/A	Unclear	High Risk	Low Risk	Low Risk	High Risk	High Risk	Unclear
Senzer, 2009	High Risk	N/A	High Risk	Unclear	Low Risk	High Risk	High Risk	High Risk	Unclear
Hu, 2006	High Risk	N/A	Unclear	Unclear	Low Risk	Unclear	High Risk	High Risk	Unclear

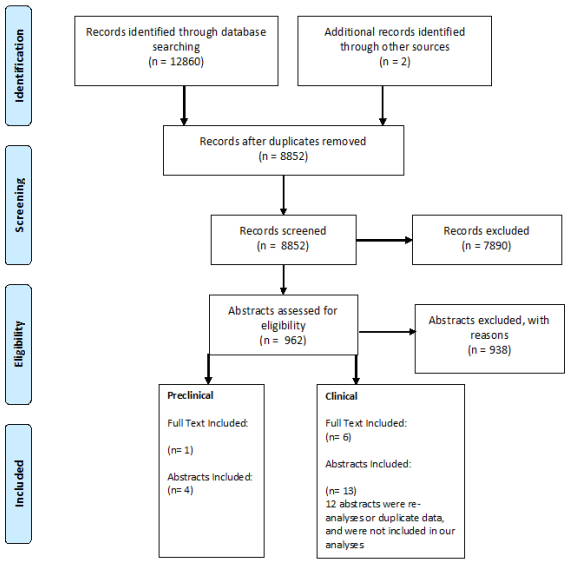

Figure 1. Study selection flow diagram

304x209mm (96 x 96 DPI)

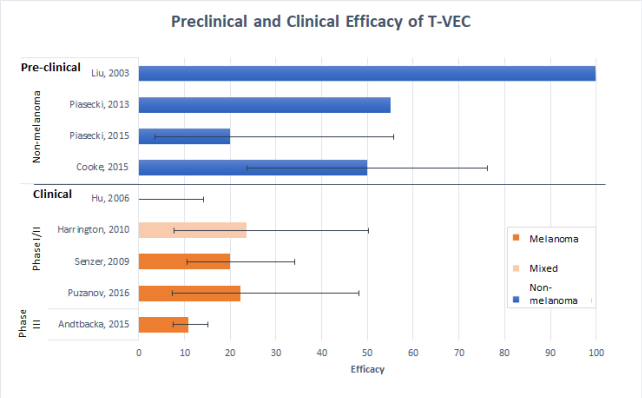

Figure 2. Preclinical and clinical efficacy of T-VEC. Four preclinical studies using mice demonstrated efficacy rates of 20-100% and clinical studies (3 melanoma and 2 mixed malignancy studies) demonstrated efficacy rates from 0-23.5%. Non-melanoma/mixed studies are represented by blue bars whereas melanoma studies are represented by orange bars. Where possible, complete regression rates of contralateral tumors for mice were used in preclinical studies and complete response was used for clinical studies. Efficacy rates decrease in the preclinical to clinical translation and upon more rigorous study design in later phase clinical trials. If possible, error bars are plotted and represent 95% confidence intervals. Some studies were not included in this analysis due to no reporting of outcome for CR.

304x209mm (96 x 96 DPI)

S1 Appendix: Search Strategy

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>

Search Strategy:

---

1 Talimogen* laherparepvec.mp. (28)

2 t vec.mp. (24)

3 OncoVEX*.mp. (15)

4 Imlygic.mp. (2)

5 JS1 34*.tw. (5)

6 or/1-5 (51)

7 Oncolytic Virotherapy/ or Oncolytic Viruses/ or cancer vaccines/tu (7016)

8 (cancer adj2 (vaccine* or virus* or virotherap* or viral therap*)).tw. (5206)

9 exp neoplasms/ or cancer.tw. (3097511)

10 or/7-9 (3097684)

11 simplexvirus/ or herpesvirus 1, human/ or Herpes Simplex/ (32955)

12 (hsv1 or hsv or herpesvirus or Herpes).tw. (72684)

13 11 or 12 (77360)

14 10 and 13 (12929)

15 ((oncolyt* or cancer or tumor or tumour) adj3 (hsv1 or hsv or hsv or herpesvirus or Herpes)).tw. (846)

16 (oncolyt* adj3 (virotherap* or virus* or viral therap*)).tw. (2183)

17 or/14-16 (14671)

18 exp animal experimentation/ or exp models, animal/ or animals/ or mammals/ or vertebrates/ or exp fishes/ or exp amphibia/ or exp reptiles/ or exp birds/ or exp hyraxes/ or exp marsupialia/ or exp monotremata/ or exp scandentia/ or exp chiroptera/ or exp carnivora/ or exp cetacea/ or exp Xenarthra/ or exp elephants/ or exp insectivora/ or exp lagomorpha/ or exp rodentia/ or exp sirenia/ or exp Perissodactyla/ or primates/ or exp strepsirhini/ or haplorhini/ or exp tarsii/ or exp platyrrhini/ or catarrhini/ or exp cercopithecidae/ or gorilla gorilla/ or pan paniscus/ or pan troglodytes/ or exp pongo/ or exp hylobatidae/ or hominidae/ (5893175)

19 (animal$1 or chordata or vertebrate* or fish$2 or amphibian* or amphibium* or reptile$1 or bird$1 or mammal* or dog or dogs or canine$1 or cat or cats or hyrax* or marsupial* or monotrem* or scandentia or bat or bats or carnivor* or cetacea or edentata* or elephant* or insect or insects or insectivore or lagomorph* or rodent$2 or mouse or mice or murine or murinae or muridae or rat or rats or pig or pigs or piglet$1 or swine or rabbit$1 or sheep$1 or goat$1 or horse$1 or equus or cow or cows or cattle or calf or calves or bovine or sirenia or ungulate$1 or primate$1 or prosimian* or haplorhini* or tarsiiform* or simian*or platyrrhini or catarrhini or cercopithecidae or ape or apes or hylobatidae or hominid* or chimpanzee* or gorilla* or orangutan* or monkey or monkeys or ape or apes).tw. (4149175)

20 (preclinic$ or pre clinic$).tw. (72274)

21 or/18-20 (6474971)

22 17 and 21 (6511)

23 6 or 22 (6545)

Supplemental Table 1. Clinical Study Characteristics

Author, Year	Country	Year Study Conducted	Study Type	Type of Cancer	Primary Outcomes	Secondary Outcomes
Andtbacka, 2015	USA, UK, Canada and South Africa	2009-2014	Interventional; Randomized (OPTiM Trial)	Melanoma	Efficacy: DRR	Efficacy: ORR OS Best Overall Response Onset and duration of response Time to treatment failure
Long 2015	USA, Australia, Switzerland, Spain	2014-2022	Interventional: Non-randomized, No control	Melanoma	Safety: Dose Limiting Toxicities	Efficacy: DRR OS Progression Free Survival Safety: AEs
Puzanov, 2015	USA	2013-2014	Interventional: Non-Randomized	Melanoma	Safety: Dose Limiting Toxicities	Efficacy: ORR Safety: Grade ≥3 AEs
Chang, 2012	USA	2006-2008	Interventional: Non-randomized, No control	Pancreatic Cancer	Efficacy: Detection of T-VEC in blood and urine Presence of Anti-HSV1 Antibodies Safety: AEs	Efficacy: ORR Change in sum of longest tumor diameter Change in pain intensity
Harrington, 2010	UK	2005-2010	Interventional: Non-randomized, No control	Squamous Cell Carcinoma	Safety: AEs	Efficacy: Antitumor Activity OS* Complete Response* Partial Response* Progression Free Survival*
Senzer 2009	USA	2005-2008	Interventional; (non-controlled, non randomized)	Melanoma	Efficacy: ORR	Efficacy: OS Safety: AEs
Hu, 2006	USA, UK	---	Interventional: Non-randomized, No control	Breast, Colorectal, Melanoma, Head and Neck	Efficacy: Biodistribution Safety: AEs	Efficacy: GM-CSF expression HSV antigen associated necrosis Viral Replication Local Reactions

---: Not Reported
*: not reported a priori
AEs - adverse events; CHN– cutaneous head and neck; DRR – durable response rate; ECOG – Eastern Cooperative Oncology Group; ORR- objective response rate; OS – overall survival

Supplemental Table 2. Clinical Patient Characteristics

Author, Year	Group	Patients (N)	Median Age (range)	Sex (n; F)	Metastasis Stage (n; Stage IVM1b/c)	Line of Therapy (n; first line)	HSV Serostatus (n; Seropositive, n; unknown)
Andtbacka, 2015	T-VEC	295	63 (22-94)	122	131	138	175, 23
Long, 2015	T-VEC + Pembrolizumab	21	58	13	11	---	---
Puzanov, 2015	T-VEC + Ipilimumab	18	---	---	---	18	---
Chang, 2012	T-VEC	17	54	6	---	---	---
Harrington, 2010	T-VEC and Chemo radiotherapy	17	58 (41-74)	2	3	---	---
Senzer, 2009	T-VEC	50	62 (34-88)	28	24	0	36, 1
Hu, 2006	T-VEC	30	55 (30-80)	23	---	0	19

--- : Not Reported

Supplemental Table 3. Clinical Intervention and Comparator Characteristics

Author, Year	Arm	Dose 1	Time of Dose 1	Frequency of Dose 1	Dose 2	Time of Dose 2	Frequency of Dose 2	Dose 3	Time of Dose 3	Frequency of Dose 3	Intervention Window	Follow Up Duration
Andtbacka, 2015	T-VEC	10^6 PFU/ml (≤4ml)	Week 1	single	10^8 PFU/ml (≤4ml)	Week 4	Q2W	N/A	N/A	N/A	Median: 23 wks (0.1-79 wks)	Median: 44 mo (32-58 mo)
	GM-CSF	125 µg/m^2	Week 1	Once daily 14/28 day cycles	N/A	N/A	N/A	N/A	N/A	N/A	Median: 10 wks (0.6 to 72 wks)	---
Long, 2015	T-VEC + Pemb.	TVEC: 10^6 PFU/ml	Day 1	single	T-VEC: 10^8 PFU/ml	Day 22	Q2W	Pemb: 200 mg	Day 36	Q2W	Median TVEC: 13 wks Median Pemb: 10 wks	---
Puzanov, 2015	T-VEC + Ipilimumab	TVEC: 10^6 PFU/ml (≤4ml)	Week 1	single	TVEC: 10^8 PFU/ml (≤4ml)	Week 4	Q2W	Ipilimumab: 3mg/kg	Week 6	Q3W	TVEC: until DLT Ipi: 12 wks	17 mo minimum
Chang, 2012	Cohort 1	10^4 PFU/ml	Week 1*	single	10^5 PFU/ml	Week 4*	Q3W*	N/A	N/A	N/A	up to 15 wks	---
	Cohort 2	10^5 PFU/ml	Week 1	single	10^6 PFU/ml	Week 4	Q3W	N/A	N/A	N/A	up to 15 wks	---
	Cohort 3	10^6 PFU/ml	Week 1	single	10^7 PFU/ml	Week 4	Q3W	N/A	N/A	N/A	up to 15 wks	---
Harrington, 2010	Cohort 1	T-VEC: 10^6 PFU/ml	Day 1	Q3W	Cisplatin: 100mg/m^2	Day 1	Q3W	N/A	N/A	N/A	Up to 9 weeks	Median:

												29mo (19-40mo)
	Cohort 2	T-VEC: 10^6 PFU/ml	Day 1	single	T-VEC: 10^7 PFU/ml	Day 22	Q3W	Cisplatin: 100mg/m^2	Day 1	Q3W	Up to 9 weeks	Median: 29mo (19-40mo)
	Cohort 3	T-VEC: 10^6 PFU/ml	Day 1	single	T-VEC: 10^8 PFU/ml	Day 22	Q3W	Cisplatin: 100mg/m^2	Day 1	Q3W	Up to 9 weeks	Median: 29mo (19-40mo)
	Cohort 4	T-VEC: 10^6 PFU/ml	Day 1	single	T-VEC: 10^8 PFU/ml	Day 22	Q3W	Cisplatin: 100mg/m^2	Day 1	Q3W	Up to 9 weeks	Median: 29mo (19-40mo)
Senzer, 2009	T-VEC	10^6 PFU/ml (≤4ml)	Week 1	Single	10^8 PFU/ml (≤4ml)	Week 4	Q2W	N/A	N/A	N/A	Max: 48 wks Median: 11 wks	Median: 18 mo (11-36 mo)
Hu, 2006	Single Dose Group 1	10^6 PFU/ml	---	single	N/A	N/A	N/A	N/A	N/A	N/A	Single dose	6 wks
	Single Dose Group 2	10^7 PFU/ml	---	single	N/A	N/A	N/A	N/A	N/A	N/A	Single dose	6 wks
	Single Dose Group 3	10^8 PFU/ml	---	single	N/A	N/A	N/A	N/A	N/A	N/A	Single dose	6 wks
	Multi-dose Group 1	10^6 PFU/ml	---	single	10^7 PFU/ml	---	Q1-3W	N/A	N/A	N/A	3-9 wks*	6 wks post final injection
	Multi-dose Group 2	10^6 PFU/ml	---	single	10^8 PFU/ml	---	Q1-3W	N/A	N/A	N/A	3-9 wks*	6 wks post final injection
	Multi-dose Group 3	10^8 PFU/ml	---	Q1-3W	N/A	N/A	N/A	N/A	N/A	N/A	3-9 wks*	6wks post final injection

DLT: Dose Limiting Toxicity; Pemb: Pembrolizumab; Q2W: every two weeks Q3W: every three weeks; Q1-3W: every 1-3 weeks; Q6W every 6 weeks
--- : not reported
T-VEC was given by intra-tumoral injection in all studies

Supplemental Table 4. Preclinical Study Characteristics

Author, Year	Year Study Conducted	Country	Study Design	Species	Strain	Model	Type of Cancer	Baseline Tumor Size	Gender	Mean Age	Mean Weight	Co-Interventions	Duration of Follow Up
Cooke, 2015	---	USA	Interventional; Non-Controlled	Mouse	Balb/c	A20 Murine Lymphoma	Lymphoma	150 mm^3	Female	---	---	N/A	---
Piasecki, 2015	---	---	Controlled Comparison	Mouse	C57Bl/6	Syngeneic MC-38	Colon Cancer	---	---	---	---	Anti-PD-1	---

						Colon Carcinoma							
Piasecki, 2013	---	---	Controlled Comparison	Mouse	---	A20 Syngeneic Contralateral Model	Lymphoma	---	---	---	---	N/A	10 days
Liu, 2003	2002	UK	Controlled Comparison	Mouse	Balb/c	Syngeneic A20 Lymphoma	Lymphoma	0.5 cm diameter	---	---	---	Immunization wild type HSV1	35 days

---: not reported
N/A: not applicable
Cooke, 2016 did not provide any relevant information

Supplemental Table 5. Preclinical Intervention and Comparator Characteristics

Author, Year	Experiment	Group	N	Dose 1	Frequency Dose 1	Duration Dose 1	Dose 2	Frequency Dose 2	Duration Dose 2	Dose 3	Frequency Dose 3	Duration Dose 3
Cooke, 2015	1	Cohort 1: TVEC	10	$3x10^4$ PFU	---	---	N/A	N/A	N/A	N/A	N/A	N/A
		Cohort 2:	10	---	---	---	N/A	N/A	N/A	N/A	N/A	N/A
		Cohort 3: TVEC	10	$3x10^6$ PFU	Every 3 days	1 week	N/A	N/A	N/A	N/A	N/A	N/A
		Cohort 4:	10	---	---	---	N/A	N/A	N/A	N/A	N/A	N/A
Piasecki, 2015	1	Int: OncoVEXmuGM-CSF + Anti-PD1	---	T-VEC: ---	Every 3 days	3 doses	Anti-PD-1: ---	Twice per wk	---	N/A	N/A	N/A
		Int: OncoVEXmuGM-CSF	---	T-VEC: ---	---	---	---	---	---	---	---	---
		Con: Anti-PD-1	---	Anti-PD-1: ---	---	---	---	---	---	---	---	---
Piasecki, 2013	1	Int: T-VEC	---	$5x10^6$ PFU	single	single	N/A	N/A	N/A	N/A	N/A	N/A
		Con: Vehicle	---	---	single	single	N/A	N/A	N/A	N/A	N/A	N/A
Liu, 2003	1	Int: JS1/34.5-/47-/mGM-CSF	10	10^6 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
			10	10^7 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
			10	10^8 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A

	Int: JS1/34.5-/47-	10	10^6 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
		10	10^7 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
		10	10^8 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
	Con: vehicle	10	50µl	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
2	Int: JS1/34.5-/47-/mGM-CSF	10	10^8 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
	Con: Vehicle	10	50µl	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A

All doses of T-VEC were given by injection intratumorally
---: not reported
Int: intervention; Con: control; wk: week; PFU: plaque forming units
Cooke, 2016 did not provide any relevant information

Supplemental Table 6. Preclinical Efficacy Data

Author, Year	Experiment	Group	N – Animals Studied	N – Lesions Studied	Baseline Mean Tumor Measure (Standard Error of Mean)	Final Mean Tumor Measure (Standard Error of Mean)	CR - Injected	CR - Contralateral	Duration of Follow Up
Cooke, 2015	1	INT: TVEC $3x10^4$ PFU	10	---	---	---	---	---	---
		INT: TVEC $3x10^6$ PFU	10	10	~150mm^3	---	10/10	5/10	---
		---	10	---	---	---	---	---	---
		---	10	---	---	---	---	---	---
Piasecki, 2015	1	INT: OncoVexmuGM-CSF	---	---	---	---	---	---	---
		INT: OncoVexmuGM-CSF + Anti-PD-1	---	20	---	---	8/10	2/10	---
		CON: Anti Pd-1	---	---	---	---	---	---	---
Piasecki, 2013	1	INT: T-VEC	---	---	---	---	70-100%	50-60%	10 days
		CON: Vehicle	---	---	---	---	---	---	---

Liu, 2003	1	INT: JS1/34.5-/47-/mGM-CSF; injected	10	N/A	5.2mm (0.34)	0.004mm (0.31)	N/A	N/A	22 days
		INT: JS1/34.5-/47-/mGM-CSF; uninjected	10 (same as injected)	N/A	5.7mm (0.29)	1.1 mm (0.73)	N/A	N/A	22 days
		INT: JS1/34.5-/47-; injected	10	N/A	5.4mm (0.37)	1.4 mm (1.36)	N/A	N/A	22 days
		INT: JS1/34.5-/47-; uninjected	10 (same as injected)	N/A	6.2mm (0.29)	5.4mm (2.01)	N/A	N/A	22 days
		CON: Vehicle; injected	10	N/A	5.4mm (0.40)	11.9mm (2.69)	N/A	N/A	22 days
		CON: Vehicle; uninjected	10 (same as injected)	N/A	5.6mm (0.46)	13.2mm (2.76)	N/A	N/A	22 days
	2	INT: JS1/34.5-/47-/mGM-CSF	10	N/A	5.5mm (0.34)	2.2mm (1.6)	N/A	N/A	21 days
		CON: Vehicle	10	N/A	5.6mm (0.23)	13.8mm (1.2)	N/A	N/A	21 days

--- : not reported
INT: intervention
CON: control
Liu 2003 data from experiment 1 taken from 10^8 dose
Cooke, 2016 did not provide any relevant information

S3. PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	4
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7

Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	7-8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	N/A

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8-9
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	14-15
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	13-14
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	14-15
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	17-18
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	18

Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	19
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	Given online

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	4
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5-6
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	5-6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	7

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	6
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	7-8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9-10
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	8-9
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	8-9
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	9-10
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11-12
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	12-13
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	13
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	14

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

BMJ Open

Mapping the preclinical to clinical evidence and development trajectory of the oncolytic virus talimogene laherparepvec (T-VEC): a systematic review

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-029475.R1
Article Type:	Original research
Date Submitted by the Author:	20-Sep-2019
Complete List of Authors:	Lalu, Manoj; Ottawa Hospital Research Institute, Clinical Epidemiology Leung, Garvin; Ottawa Hospital Research Institute, Clinical Epidemiology Dong, Yuan Yi; Ottawa Hospital Research Institute, Clinical Epidemiology Montroy, Joshua; Ottawa Hospital Research Institute, Clinical Epidemiology Butler, Claire; Ottawa Hospital Research Institute, Clinical Epidemiology Auer, Rebecca; Ottawa Hospital Research Institute, Clinical Epidemiology Fergusson, Dean; Ottawa Hospital Research Institute, Medicine
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Oncology
Keywords:	TVEC, oncolytic virus, cancer, translation, review

Mapping the preclinical to clinical evidence and development trajectory of the oncolytic virus talimogene laherparepvec (T-VEC): a systematic review

Manoj M. Lalu† MD, PhD, FRCPC[1,2,3,4], Garvin J. Leung† MPH[2,5], Yuan Yi Dong[2,5], Joshua Montroy MSc,[2] Claire Butler[2], Rebecca C. Auer MD, MSc, FRCSC[7,8], Dean A. Fergusson PhD, MHA*[2,6,7]

[1]Department of Anesthesiology and Pain Medicine, The Ottawa Hospital, University of Ottawa, Ottawa, ON, Canada
[2]Blueprint Translational Research Group, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
[3] Regenerative Medicine Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
[4] Department of Cellular and Molecular Medicine, University of Ottawa , Ottawa, ON, Canada
[5]Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada
[6]School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada
[7]Department of Surgery, The Ottawa Hospital, University of Ottawa, Ottawa, ON, Canada
[8]Cancer Therapeutics Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
† These authors contributed equally to this work

Emails: MML: mlalu@toh.ca; GJL: garvin.leung@uottawa.ca; YYD: ydong044@uottawa.ca; JM: jmontroy@ohri.ca; CB: clbutler@toh.ca; RCA: rauer@toh.ca; DAF: dafergusson@ohri.ca

*Corresponding author:
Dean A. Fergusson email: dafergusson@ohri.ca
Centre for Practice-Changing Research,
Office L1298a
501 Smyth Road, Box 201B
Ottawa, Ontario, Canada K1H 8L6
Tel. 1-613-737-8480
Fax. 1-613-739-6938

Running Head: The efficacy of T-VEC: A systematic review

Word count: 2631

ABSTRACT

Objective: This study aimed to conducta systematic review of preclinical and clinical evidence to chart the successful trajectory of Talimogene laherparepvec (T-VEC), from the bench to the clinic.

Design: This study was a systematic review. The primary outcome of interest was the efficacy of treatment, determined by complete response. Abstract and full-text selection as well as data extraction was done by two independent reviewers. The Cochrane risk of bias tool was used to assess the risk of bias in studies.

Setting: Embase, Embase Classic, and OvidMedline were searched from inception until May 2016 to assess its development trajectory to approval in 2015.

Participants: Preclinical and clinical controlled comparison studies, as well as observational studies.

Interventions: T-VEC for treatment of any malignancy.

Results: 8,852 records were screened and five preclinical (n=150 animals) and seven clinical studies (n=589 patients) were included. We saw large decreases in T-VEC's efficacy as studies moved from the laboratory to patients, and as studies became more methodologically rigorous. Preclinical studies reported complete regression rates up to 100% for injected tumors and 80% for contralateral tumors, while the highest degree of efficacy seen in the clinical setting was a 24% complete response rate, with one study experiencing a complete response rate of 0%. We were unable to reliably assess safety due to the lack of reporting, as well as the heterogeneity seen in adverse event definitions. All preclinical studies had high or unclear risk of bias, and all clinical studies were at a high risk of bias in at least one domain.

Conclusions: Our findings illustrate that even successful biotherapeutics may not demonstrate a clear translational road map. This emphasizes the need to consider increasing rigour and transparency along the translational pathway.

PROSPERO Registration Number: CRD42016043541

Keywords: T-VEC, oncolytic virus, cancer, translation, review

Strengths and limitations of this study

- Comprehensive, up-to-date review of the efficacy and safety of T-VEC
- Threats to both internal validity and construct validity were performed
- Reporting of methods and findings was incomplete in most of the studies included
- Poor reporting and study design are major contributors to the ongoing reproducibility crisis in preclinical research

BACKGROUND

Preclinical research receives approximately half of the world's biomedical research funding, yet very few of its findings translate clinically. This represents an enormous waste of resources with an estimated 28 billion dollars per year in the US alone being spent on biomedical research which is not reproducible and therefore not translatable.(1) One study found that only 5% of highly efficacious preclinical therapeutics were clinically translated.(2) These successes often take almost twenty years to become successfully translated across the research spectrum. (2, 3)

Although the process of clinical translation is complicated, the transition from bench-to-bedside often starts with preclinical research. These investigations (usually on animals or cells), are aimed at studying efficacy, pharmacokinetics and dynamics, as well as detailing safety.(4) Next, a drug is tested in a phase I clinical trial, which usually contains a small number of participants and is aimed at studying the safety of the drug. If a drug is safe, it may proceed to phase II which are larger than phase I studies and are designed to test safety, pharmacokinetics, pharmacodynamics, and optimal dosing regimens. They may also offer preliminary evidence of drug efficacy. Finally, a methodologically rigorous phase III study is performed. These studies are designed and powered to test efficacy in the patient population of interest (usually against a comparator such as placebo), as well as identify rarer adverse events which may have gone unnoticed in a smaller phase I or II study.(5)

Given the high failure rate in translating therapies across this spectrum, as well as significant time-lags associated with translation, it is important that we examine the few agents that have successfully crossed the preclinical-to-clinical bridge in order to learn from and replicate their success. Thus, we conducted a comprehensive evaluation of available evidence supporting the successful translation of Talimogene laherparepvec (T-VEC). T-VEC is a modified HSV-1 virus

produced by *Amgen* and it is the first, and only, FDA approved oncolytic virus therapy; it is currently approved to treat advanced melanoma.(6) Oncolytic viruses are an emerging cancer therapy that work by preferentially targeting and infecting cancer cells.(6) Upon infection, oncolytic viruses can induce an anti-tumor immune response that reduces tumor burden.

Through a careful evaluation of T-VEC development we hoped to identify factors that may contribute to bench-to-bedside success. This may serve an exemplar for other therapies as they move along the translational continuum. Thus, the purpose of this systematic review was to map the successful preclinical to clinical trajectory of T-VEC to inform the development paths of future biotherapeutics.

METHODS

Our review was registered in full on PROSPERO, the international prospective register of systematic reviews (no. CRD42016043541). The review is reported in accordance to the PRISMA guidelines.(7)

Eligibility Criteria

We included all clinical and preclinical *in vivo* controlled comparison studies of T-VEC for treatment of any malignancy (randomized, pseudo-randomized, and non-randomized studies), as well as observational studies such as case-control, case-series and case reports. Studies reporting only *ex vivo* or *in vitro* experiments were excluded. For both preclinical and clinical studies, we included studies that administered T-VEC as a monotherapy or in combination with other therapies for treatment of malignancy. We had no exclusions on comparison treatments, which include standard line therapy or no treatment.

Outcomes

The primary outcome of interest was the efficacy of treatment. Our primary indicator of efficacy was complete response. Other measures of efficacy such as survival, response rates (durable, partial, objective), time to treatment failure, and disease stability were also collected. Such measures were based on the Response Evaluation Criteria in Solid Tumors (RECIST) Guidelines.(8) In preclinical studies, additional measures of efficacy such as changes in mean tumor volume and number of lesions were collected. The primary indicator of efficacy, complete response, was used as the primary outcome regardless of reporting within the individual study, in order to assess the continuity of evidence along the research spectrum. The secondary outcome of

interest was safety, for which we collected data on all adverse events in preclinical and clinical studies.

Literature Search

In collaboration with a medical information specialist (Risa Shorr, Learning Services, The Ottawa Hospital) a search strategy was designed to identify all relevant preclinical and clinical studies. Searches were conducted in the following databases: Embase, Embase Classic, and OvidMedline from inception until May 2016. This time frame was chosen to ensure all published studies that contributed T-VECs FDA approval in 2015 were included. Search terms included: Talimogen laherparepvec, Tvec, OncoVEX and Imlygic. Additional terms pertaining to preclinical studies (e.g. animal experiment/model) and oncology (e.g. cancer, neoplasm, oncolytic virus) were also included. Studies were also screened for inclusion based on reference tracking, by scanning the bibliography of included primary studies and relevant review articles. We did not impose any restrictions on language or publication type. A grey literature search was not performed. The finalized search strategy can be found in online supplementary file 1.

Study Selection Process

Studies identified by our literature search were collated and duplicates were removed. Titles and abstracts were independently screened for inclusion by two reviewers using DistillerSR (Evidence Partners, Ottawa, ON). Those deemed potentially relevant were recorded, and full-text articles were obtained. The same reviewers screened full articles for final eligibility. Disagreements at any stage were resolved by discussion or by consultation with a senior team member when necessary. The study selection process was documented using a PRISMA flow diagram (Figure 1).

Data Extraction

All data extraction was completed independently and in duplicate, using a standardized and piloted data extraction form, with disagreements resolved as mentioned above. Data pertaining to general and intervention characteristics of the included studies were extracted (e.g. study design, country, type of malignancy, dosing of intervention and comparator treatments). For clinical studies, data was collected on patient characteristics (e.g. age, sex, cancer staging, HSV status). For preclinical studies, characteristics on the animal model were extracted (e.g. type of species, cell line used, disease induction method, age, sex, weight).

Risk of bias – assessment to risk of internal validity

Clinical studies that met inclusion criteria were assessed for risk of bias in duplicate, according to the recommended methodology of the Cochrane Collaboration.(9) Five types of biases (selection, performance, detection, attrition, reporting biases) were assessed using six domains: randomization, allocation concealment, blinding of participants/personnel, outcome assessment blinding, incomplete outcome reporting, and selective outcome reporting. Additional domains assessed for risk of bias were: i) reported conflicts of interest, ii.) sample size calculation, and iii.) funding. Each domain was given a score of "high", "unclear", or "low" risk of bias for each study. Risk of bias assessment for preclinical studies were assessed using a modified Cochrane Risk of Bias tool and assessed the same domains as indicated for clinical studies.(10)

Assessment of threats to construct validity

Construct validity is the concept in how much a preclinical experiment (i.e. animal studies) corresponds to the clinical entity it is intended to model. There are various threats to construct validity that can be introduced from the preclinical study design. The items evaluated in duplicate

for each preclinical study include: i.) use of adult animals, ii.) use of animals with advanced stage disease (defined as the presence of multiple visceral lesions and/or clinical/histological signs of malignant progression), iii.) immune status of animals to HSV, iv.) whether a xenograft model was used, and v.) the use of a humanized immune system model. Each of these items was given a score of "yes", "no", or "unclear" for every preclinical study.

Statistical Analysis

Efficacy was expressed as proportions with accompanying 95% confidence intervals. If confidence intervals were not present within the individual study, they were calculated via standard methods.(11) To assess the continuity between preclinical and clinical studies, the efficacy of studies was plotted as percentage response.

Deviations from Protocol

We were unable to assess safety as we could not acquire patient-level safety data. Furthermore, our primary efficacy outcome stated in protocol was durable response rate. However, this was changed to complete response as most clinical studies did not report durable response. Subgroup analyses, meta-analyses, Egger's test, and pooling of data could not be conducted due to the limited available data.

Patient and Public Involvement

Patients and the public were not involved in this research.

RESULTS

Upon removal of duplicates, a total of 8,852 references were identified by the electronic search. During the review of titles and abstracts, 7,890 references were excluded. Following full text

screening, a total of 7 clinical studies,(12-18) and 5 preclinical studies(19-23) were included in our review (Figure 1).

Characteristics of Included Trials

Characteristics of included studies are shown in Table 1. Preclinical studies were published between 2003 and 2016 and sample sizes ranged from 20 to 90. Of the 5 preclinical studies, three used a lymphoma model, one used a colorectal model, and one used a melanoma model. All studies were performed in mice. The duration of follow-up was reported by two studies and ranged from 10 days to 35 days. The dose of T-VEC used ranged from $3x10^4$ plaque forming units (PFU) to $5x10^6$ PFU. One frequency of T-VEC administration varied from, every three days for one week, every three days for nine days, a single dose given only once, and every other day for five days. Specific details of study and intervention characteristics for each preclinical study can be found in online supplementary files 2 and 3.

Clinical studies were published between 2006 and 2016 and took place in seven countries. Sample sizes ranged from 17 to 295. Of the seven clinical studies, four were in melanoma patients, one was in pancreatic cancer patients, one in head and neck cancer patients and one studied a mix of breast, colorectal, melanoma and head and neck cancer patients. Six were either Phase I or II, and one trial was a Phase III evaluation. The primary outcome was efficacy in two studies, safety in three studies and a combination of efficacy and safety in the other two studies. The duration of follow-up ranged from six weeks to 44 months.

T-VEC was administered alone in four studies, while it was administered adjuvant to chemotherapy in 3 studies. The dose of T-VEC administered ranged from 10^4 PFU/mL to 10^8 PFU/mL. In the large, Phase III study, T-VEC was administered at ≤4mL x 10^6 PFU/mL once, and then three weeks later, ≤4mL 10^8 PFU/mL was administered every two weeks for a median of 23

weeks. A similar dosing regimen was used in three other trials. The other three trials were dose-finding in nature and had multiple trial arms receiving increasing doses of T-VEC. In-depth study details, as well as participant and intervention details for each study can be found in online supplementary files 4, 5, and 6.

Efficacy of Treatment

Treatment efficacy for each study is summarized in Table 1 and Figure 2. Preclinical studies reported complete regression rates up to 100% for injected tumors and 80% for contralateral tumors (see also online supplementary file 7). In comparison, the first published Phase I T-VEC clinical trial reported a complete response of 0% for cutaneous lesions caused by malignancies of head and neck, breast, colorectal, and melanoma. Of the multiple malignancies treated, melanoma had the best response in this trial. Subsequent Phase I/II melanoma trials were then conducted and demonstrated complete response rates of 20-22%. This was followed by the Phase III OPTIM melanoma trial, which had a complete response rate of 10.8%. Studies involving non-melanoma cancers varied with efficacies between 0-24%

Safety of treatment

We attempted to assess safety across clinical studies, however we were unable to obtain patient level data from any of the studies. The definitions of adverse events, and the manner in which they were classified, was found to be highly heterogenous across studies. Studies did not specify what percent of adverse events were repeated adverse events from the same patient(s), used different criteria for recording and reporting adverse events, categorized them differently, etc. Therefore, we were unable to pool adverse events or interpret findings reliably.

Validity Assessments

Construct validity, the concept of how well an animal model represents the clinical entity it is intended to mimic, was first assessed through the following domains: the use of appropriately-aged mice, advanced stage of disease, HSV-immunity, and types of mouse models. None of the preclinical studies fully reported or used methodologies to reduce threats to construct validity domains (Table 2). No studies declared using adult animal models, no studies used animals with late stage disease, only one study used animals immune to HSV, no studies used a xenograft model, and no studies reported using an animal model with a humanized immune system.

We also assessed internal validity (i.e. risk of bias) and found that all preclinical studies had high or unclear risk of bias across the assessed domains: randomization sequence, allocation concealment, blinding, incomplete reporting, sample size calculation, and funding source (Table 3). For clinical studies, early phase trials had high or unclear risk of bias across at least six of nine domains whereas the more robust Phase III OPTIM trial had the lowest risk of bias and also the lowest efficacy of any of the published melanoma clinical trials (Table 4). Reporting of key methodological elements was lacking.

DISCUSSION

We hoped to synthesize the evidence to produce a clear road map of T-VEC's translation in the published literature to follow the journey of a successful biotherapeutic. Yet, we were unable to paint a clear picture of how the evidence was utilized in proceeding to melanoma clinical trials. Rather, our assessment uncovered a disconnect between *in vivo* preclinical and clinical findings. Furthermore, the road map was plagued with poor reporting, high risk of bias, and insufficient data along the translational path. Overall, we were surprised by the pace and magnitude of diminishing

efficacy as T-VEC moved from bench to bedside and then towards later phase clinical trials (i.e. Phase I to III). Although T-VEC was successful in terms of gaining regulatory approval, its translational path is complicated, and the pieces of the evidence puzzle do not easily fit together. While we appreciate that translation is not a predictable linear process, it is difficult to learn from the example of T-VEC given the available and reported pre-clinical and clinical evidence.

While many novel therapeutics are under intellectual property rights, details of study design and results should be transparently reported for scientists, clinicians, and patients to evaluate findings. The fact that the only FDA approved oncolytic virus therapy is not clearly reported illustrates the issues plaguing the success of cancer therapeutics. Nonetheless, T-VEC has shown some efficacy in treating refractory melanoma and numerous clinical trials are underway to assess its use in combination with other cancer regimens and in treating other malignancies.

Perhaps the largest discrepancy noted was that only a single preclinical study used a melanoma model, whereas all but two clinical studies administered T-VEC to melanoma patients. Conversely, lymphoma, which was used in three preclinical studies, was not assessed in clinical studies. Interestingly, our subsequent searches found that Amgen's FDA filing (STN# 125518.000) (24) for T-VEC did not appear to report on any *in vivo* melanoma models, whereas the EMA report did (EMA/734400/2015).(25) Thus, the majority of animal models were off-target from the malignancies studied in clinical trials and may have poorly represented melanoma in the clinical setting. Coupled with these findings was the fact that the majority of our studies were found to be at a high risk of bias.Such threats to internal validity can bias results and may help explain T-VEC's superior preclinical efficacy compared to later phase clinical trials. A lack of randomization and blinding in preclinical studies has been associated with inflated effect sizes,(26, 27) thus this may partially explain the preclinical to clinical discrepancy of T-VEC.

Reporting of methods and findings was incomplete in most of the studies included. Only one full preclinical article on T-VEC was published, and solely aggregate patient data for later phase trials was available. Poor reporting and study design are major contributors to the ongoing reproducibility crisis in preclinical research.(28) Thus, in hopes of presenting a clearer picture of T-VEC's successful translation, we contacted *Amgen* to obtain preclinical *in vivo* melanoma data, patient-level safety data, and any additional efficacy data. Patient-level data would afford the ability to combine data across T-VEC's clinical development and also provide clarification into the categorization of adverse events. Recently, release of individual patient data to third parties has been advocated by the Institute of Medicine, journal editors, and others as it enhances transparency, enables re-analyses of data, and helps address reproducibility.(29) The reporting of harms in clinical trials remains an issue in the scientific community,(30-32) and represents a roadblock to translational success. Some basic steps required to improve the reporting of safety in translational research include the development of standardized scales and instruments, instituting active rather than passive surveillance for toxicity, including detailed information on participant withdrawals due to toxicity, reporting the timing, frequency, and duration of clinically relevant events, and the publication of raw data.(33, 34) *Amgen,* however, was unwilling to enter a data sharing agreement, as they stated that there was little value to compel a transparent data release for our proposed analyses. This lack of transparency and incomplete reporting is disappointing, especially considering that it was *Amgen* that previously highlighted poor reporting as contributing to its own failure to reproduce 47 of 53 high-impact preclinical cancer studies.(35) Their findings fuelled a call by the NIH and other stakeholders to enhance the reproducibility and transparency of preclinical research.(36)

As stated, we recognize that translation is not a linear process, but we should observe consistent and coherent patterns. Moving forward, we suggest that preclinical and clinical studies for emerging therapies should be fully reported and attention should be given to validities in order to develop more precise estimates of effect early in development. Investigators should carefully match their preclinical model to the intended clinical population; when possible, both disease states and outcomes measured should have high construct validity. Following successful exploratory preclinical studies, investigators should consider preclinical systematic reviews(37) and designing methodologically rigorous confirmatory and/or multicenter preclinical studies.(38) These steps may allow preclinical testing to more accurately forecast downstream clinical results in human patients.(27) Within the trajectory of clinical development (i.e. once clinical trials have been initiated), careful consideration of methods to reduce bias should also be considered (although, this may not be possible for the earliest phase trials). We believe these steps will provide unbiased and valuable information that will ultimately provide patients with cancer therapies that match their preclinical and early clinical promise.

CONCLUSIONS

The findings from our systematic review demonstrate that even successful biotherapeutics may not demonstrate a clear translational road map. The magnitude of efficacy of T-VEC demonstrated in preclinical studies was considerably larger when T-VEC was moved to the clinic, and the most methodologically rigorous trial included in our review demonstrated the smallest degree of efficacy. Methodologically rigorous studies should be performed earlier on in the translational pathway, which may help to get a realistic estimate of treatment efficacy prior to clinical translation.

DECLARATIONS

Ethics approval: Not applicable.

Consent for publication: Not applicable.

Availability of data: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding: Biotherapeutics for Cancer Treatment (BioCanRx) supported the conduct of this study by a Catalyst grant. BioCanRx is a Government of Canada funded Networks of Centres of Excellence and was not involved in any other aspect of the project, such as the design of the project's protocol and analysis plan, the collection of data and analyses. CB was also supported by a BioCanRx studentship. ML is supported by The Ottawa Hospital Anesthesia Alternate Funds Association and the Scholarship Protected Time Program, Department of Anesthesiology and Pain Medicine, uOttawa.

Author contributions: MML and DAF conceptualized the study. MML, RA, DAF, and GJL contributed to the study design. GJL, YYD, and CB conducted data extraction. All authors analysed and interpreted the data. MML and GJL were responsible for drafting the manuscript. All authors critically reviewed the manuscript and provided intellectual content. All authors approve the final version of the manuscript.

Acknowledgments: We thank Risa Shorr (Ottawa Hospital Research Institute librarian) for systematic search assistance. MML is supported by The Ottawa Hospital Anesthesia Alternate Funds Association and the Scholarship Protected Time Program, Department of Anesthesiology and Pain Medicine, uOttawa.

REFERENCES

1. Freedman LP, Cockburn IM, Simcoe TS. The Economics of Reproducibility in Preclinical Research. PLoS Biol. 2015;13(6):e1002165.
2. Contopoulos-Ioannidis DG, Ntzani E, Ioannidis JP. Translation of highly promising basic science research into clinical applications. Am J Med. 2003;114(6):477-84.
3. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. J R Soc Med. 2011;104(12):510-20.
4. Use of Laboratory Animals in Biomedical and Behavioral Research. Washington (DC)1988.
5. Umscheid CA, Margolis DJ, Grossman CE. Key concepts of clinical trials: a narrative review. Postgrad Med. 2011;123(5):194-204.
6. Rehman H, Silk AW, Kane MP, Kaufman HL. Into the clinic: Talimogene laherparepvec (T-VEC), a first-in-class intratumoral oncolytic viral therapy. J Immunother Cancer. 2016;4:53.
7. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009;62(10):1006-12.
8. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47.
9. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.
10. Hooijmans CR, Rovers MM, de Vries RBM, Leenaars M, Ritskes-Hoitinga M, Langendam MW. SYRCLE's risk of bias tool for animal studies. BMC Medical Research Methodology. 2014;14(1):43.
11. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Statistics in medicine. 1998;17(8):857-72.
12. Hu JC, Coffin RS, Davis CJ, Graham NJ, Groves N, Guest PJ, et al. A phase I study of OncoVEXGM-CSF, a second-generation oncolytic herpes simplex virus expressing granulocyte macrophage colony-stimulating factor. Clinical cancer research : an official journal of the American Association for Cancer Research. 2006;12(22):6737-47.
13. Senzer NN, Kaufman HL, Amatruda T, Nemunaitis M, Reid T, Daniels G, et al. Phase II clinical trial of a granulocyte-macrophage colony-stimulating factor-encoding, second-generation oncolytic herpesvirus in patients with unresectable metastatic melanoma. J Clin Oncol. 2009;27(34):5763-71.
14. Andtbacka RHI, Kaufman HL, Collichio F, Amatruda T, Senzer N, Chesney J, et al. Talimogene Laherparepvec Improves Durable Response Rate in Patients With Advanced Melanoma. Journal of Clinical Oncology. 2015;33(25):2780-8.
15. Long GV, Dummer R, Ribas A, Puzanov I, Michielin O, Vanderwalde A, et al. 24LBA Safety data from the phase 1b part of the MASTERKEY-265 study combining talimogene laherparepvec (T-VEC) and pembrolizumab for unresectable stage IIIB-IV melanoma. European Journal of Cancer. 2015;51:S722.
16. Puzanov I, Milhem MM, Minor D, Hamid O, Li A, Chen L, et al. Talimogene Laherparepvec in Combination With Ipilimumab in Previously Untreated, Unresectable Stage IIIB-IV Melanoma. J Clin Oncol. 2016;34(22):2619-26.

17. Harrington KJ, Hingorani M, Tanay MA, Hickey J, Bhide SA, Clarke PM, et al. Phase I/II study of oncolytic HSV GM-CSF in combination with radiotherapy and cisplatin in untreated stage III/IV squamous cell cancer of the head and neck. Clinical cancer research : an official journal of the American Association for Cancer Research. 2010;16(15):4005-15.

18. Chang KJ, Senzer NN, Binmoeller K, Goldsweig H, Coffin R. Phase I dose-escalation study of talimogene laherparepvec (T-VEC) for advanced pancreatic cancer (ca). Journal of Clinical Oncology. 2012;30(15_suppl):e14546-e.

19. Liu BL, Robinson M, Han ZQ, Branston RH, English C, Reay P, et al. ICP34.5 deleted herpes simplex virus with enhanced oncolytic, immune stimulating, and anti-tumour properties. Gene Ther. 2003;10(4):292-303.

20. Piasecki J, Tiep L, Zhou J, Beers C. Talilmogene Iaherparepvec generates systemic T-cell-mediated anti-tumor immunity. Journal for ImmunoTherapy of Cancer. 2013;1(1):P198.

21. Cooke K, Rottman J, Zhan J, Mitchell P, Ikotun O, Yerby B, et al. Oncovex MGM-CSF –mediated regression of contralateral (non-injected) tumors in the A20 murine lymphoma model does not involve direct viral oncolysis. Journal for Immunotherapy of Cancer. 2015;3(Suppl 2):P336-P.

22. Piasecki J, le T, Ponce R, Beers C. Abstract 258: Talilmogene laherparepvec increases the anti-tumor efficacy of the anti-PD-1 immune checkpoint blockade. Cancer Research. 2015;75(15 Supplement):258-.

23. Cooke K, Estrada J, Zhan J, Mitchell P, Bulliard Y, Beltran PJ. Abstract 2351: Development of a B16F10 cell line expressing mNectin1 to study the activity of OncoVEXmGM-CSF in murine syngeneic melanoma models. Cancer Research. 2016;76(14 Supplement):2351-.

24. US Food and Drug Administration. IMLYGIC (talimogene laherparepvec) 2015 [Available from: https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/imlygic-talimogene-laherparepvec.

25. European Medicines Agency. Assessment report Imlygic International non-proprietary name: talimogene laherparepvec 2015 [Available from: https://www.ema.europa.eu/en/documents/assessment-report/imlygic-epar-public-assessment-report_en.pdf.

26. Hirst JA, Howick J, Aronson JK, Roberts N, Perera R, Koshiaris C, et al. The need for randomization in animal trials: an overview of systematic reviews. PLoS One. 2014;9(6):e98856.

27. Macleod MR, van der Worp HB, Sena ES, Howells DW, Dirnagl U, Donnan GA. Evidence for the efficacy of NXY-059 in experimental focal cerebral ischaemia is confounded by study quality. Stroke. 2008;39(10):2824-9.

28. Landis SC, Amara SG, Asadullah K, Austin CP, Blumenstein R, Bradley EW, et al. A call for transparent reporting to optimize the predictive value of preclinical research. Nature. 2012;490(7419):187-91.

29. Zarin DA, Tse T. Sharing Individual Participant Data (IPD) within the Context of the Trial Reporting System (TRS). PLoS Med. 2016;13(1):e1001946.

30. Hodkinson A, Kirkham JJ, Tudur-Smith C, Gamble C. Reporting of harms data in RCTs: a systematic review of empirical assessments against the CONSORT harms extension. BMJ Open. 2013;3(9):e003436.

31. Rodgers MA, Brown JV, Heirs MK, Higgins JP, Mannion RJ, Simmonds MC, et al. Reporting of industry funded study outcome data: comparison of confidential and published data on the safety and effectiveness of rhBMP-2 for spinal fusion. BMJ. 2013;346:f3981.

32. Jureidini JN, McHenry LB, Mansfield PR. Clinical trials and drug promotion: selective reporting of study 329. International Journal of Risk & Safety in Medicine. 2008;20(1-2):73-81.

33. Ioannidis JP, Lau J. Improving safety reporting from randomised trials. Drug Saf. 2002;25(2):77-84.

34. Lineberry N, Berlin JA, Mansi B, Glasser S, Berkwits M, Klem C, et al. Recommendations to improve adverse event reporting in clinical trial publications: a joint pharmaceutical industry/journal editor perspective. BMJ. 2016;355:i5078.

35. Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. Nature. 2012;483(7391):531-3.

36. Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. Nature. 2014;505(7485):612-3.

37. Lalu MM, Sullivan KJ, Mei SH, Moher D, Straus A, Fergusson DA, et al. Evaluating mesenchymal stem cell therapy for sepsis with preclinical meta-analyses prior to initiating a first-in-human trial. Elife. 2016;5.

38. Llovera G, Hofmann K, Roth S, Salas-Perdomo A, Ferrer-Ferrer M, Perego C, et al. Results of a preclinical randomized controlled multicenter trial (pRCT): Anti-CD49d treatment for acute brain ischemia. Science translational medicine. 2015;7(299):299ra121.

Figure Legends

Figure 1. Study selection flow diagram

Figure 2. Preclinical and clinical efficacy of T-VEC. Four preclinical studies using mice demonstrated efficacy rates of 20-100% and clinical studies (3 melanoma and 2 mixed malignancy studies) demonstrated efficacy rates from 0-23.5%. Non-melanoma/mixed studies are represented by blue bars whereas melanoma studies are represented by orange bars. Where possible, complete regression rates of contralateral tumors for mice were used in preclinical studies and complete response was used for clinical studies. Efficacy rates decrease in the preclinical to clinical translation and upon more rigorous study design in later phase clinical trials. If possible, error bars are plotted and represent 95% confidence intervals. Some studies were not included in this analysis due to no reporting of outcome for CR.

Table 1A. *Study characteristics of included preclinical studies of T-VEC.*

Preclinical Study	Treatment	Total Number of Animals Used	Type of Cancer/Model	Efficacy Measures*	Risk of Bias (/9**)
Liu, 2003[14]	T-VEC; HSV1 wildtype immunization	90	Lymphoma (A20 murine lymphoma mouse model)	CR: 100% (n=10) (injected)	9
Piasecki, 2013[15]	T-VEC	NR	Lymphoma (A20 murine lymphoma mouse model)	CR: 70-100% of injected, 50-60% of contralateral	9
Piasecki, 2015[17]	T-VEC + Anti-PD-1	NR	Colorectal (MC-38 colon carcinoma mouse model)	CR: 80.0% (44.2-96.5%) (injected) n=10 CR: 20.0% (3.5-55.8%) (contralateral) n=10	9
Cooke, 2015[16]	T-VEC	40	Lymphoma (A20 murine lymphoma mouse model)	CR: 100% (65.5-100%) (injected) n=10 CR: 50% (23.7-76.3%) (contralateral)	9
Cooke, 2016[18]	T-VEC	20	Melanoma (B16F10 melanoma model)	NR – statistically significant tumor reduction and survival noted.	9

Table 2B. *Study characteristics of included clinical studies of T-VEC.*

Clinical Study	Treatment	Total N	Type of Cancer	Efficacy Measures*	Risk of Bias (/9**)
Hu, 2006[7] *Non-controlled Phase I*	T-VEC	30 (9 melanoma)	Breast, colorectal, melanoma, head and neck	CR: 0% (0-14.1%) PR: 0% (0-14.1%)	7
Senzer, 2009[8] *Non-controlled Phase II*	T-VEC	50	Melanoma	OR: 26.0% (15.1-40.6%) CR: 20.0% (10.5-34.1%) PR: 10.0% (3.7-22.6%)	7
Harrrington, 2010[12] *Phase I/II*	T-VEC + cisplatin	17	Head and neck	CR: 23.5% (7.8-50.2%) PR: 58.8% (33.5-80.6%) OR: 82.4% (55.8-95.3%)	6
Chang, 2012[13] *Phase I*	T-VEC	17	Pancreatic	OR: 0% (0-22.9%)	6
Andtbacka, 2015[9] *Phase III*	T-VEC	295	Melanoma	DR: 16.3% (12.1-20.5%) OR: 26.4% (21.4-31.5%) PR: 15.6% (11.7-20.3%) CR: 10.8% (7.6-15.1%)	3
	GM-CSF (control)	141		DR: 2.1% (0-4.5%) OR: 5.7% (1.9-9.5%) PR: 5.0% (1.3-8.5%) CR: < 1%	
Long, 2015[10] *Phase Ib*	T-VEC + pembrolizumab	21	Melanoma	-	6
Puzanov, 2016[11] *Phase Ib*	T-VEC + IPI	18	Melanoma	DR: 44.4% (22.4-68.7%) OR: 50.0% (29.0-70.9%) CR: 22.2% (7.4-48.1%) PR: 27.8% (10.7-53.6%)	6

* DR – durable response; OR – objective response; CR – complete response/complete regression; PR – partial response; DR/OR/CR/PR definitions were based on RECIST guidelines for clinical studies. **Total number of domains that were assessed a score of high risk or unclear (maximum = 9). The nine domains include randomization, allocation concealment, blinding of participants/personnel, outcome assessment blinding, incomplete outcome reporting, selective outcome reporting, reported conflicts of interest, sample size calculation, funding.

Table 3. Construct validity assessment for preclinical studies

Author, Year	Adult Used	Animals with Advanced Stage Disease	Animals Immune to HSV	Xenograft Model Used	Used Model with a Humanized Immune System
Cooke, 2016	Unclear	No	Unclear	No	Unclear
Cooke, 2015	Unclear	No	Unclear	No	Unclear
Piasecki, 2015	Unclear	Unclear	Unclear	No	Unclear
Piasecki, 2013	Unclear	Unclear	Unclear	No	Unclear
Liu, 2003	Unclear	No	Yes	No	Unclear

Table 4. Risk of bias assessment for preclinical studies

Author, Year	Random Sequence Generation	Allocation Concealment	Blinding of Personnel	Blinded Outcome Assessment	Incomplete Outcomes Addressed	Selective Outcome Reporting	Conflicts of Interest	A Priori Sample Size Calculation	Funding
Cooke, 2016	High Risk	High Risk	Unclear	Unclear	High Risk	High Risk	Unclear	Unclear	High Risk
Cooke, 2015	High Risk	High Risk	Unclear	Unclear	High Risk	High Risk	Unclear	Unclear	High Risk
Piasecki, 2015	Unclear	Unclear	Unclear	Unclear	High Risk	High Risk	Unclear	Unclear	High Risk
Piasecki, 2013	Unclear	Unclear	Unclear	Unclear	High Risk	High Risk	Unclear	Unclear	High Risk
Liu, 2003	Unclear	Unclear	Unclear	Unclear	High Risk	Unclear	Unclear	Unclear	High Risk

Table 5. Risk of bias assessment for clinical studies

Author, Year	Random Sequence Generation	Allocation Concealment	Blinding of Participants and Personnel	Blinding of Outcome Assessors	Incomplete Outcome Data Addressed	Selective Reporting	Conflicts of Interest	Funding	Sample Size Calculation
Andtbacka, 2015	Low Risk	Low Risk	High Risk	Low Risk	Low Risk	Low Risk	High Risk	High Risk	Low Risk
Long, 2015	High Risk	N/A	High Risk	Unclear	Low Risk	Low Risk	High Risk	High Risk	Unclear
Puzanov, 2016	High Risk	N/A	High Risk	Unclear	Low Risk	Low Risk	High Risk	High Risk	Unclear
Chang, 2012	High Risk	N/A	High Risk	Unclear	Low Risk	Low Risk	High Risk	High Risk	Unclear
Harrington, 2010	High Risk	N/A	Unclear	High Risk	Low Risk	Low Risk	High Risk	High Risk	Unclear
Senzer, 2009	High Risk	N/A	High Risk	Unclear	Low Risk	High Risk	High Risk	High Risk	Unclear
Hu, 2006	High Risk	N/A	Unclear	Unclear	Low Risk	Unclear	High Risk	High Risk	Unclear

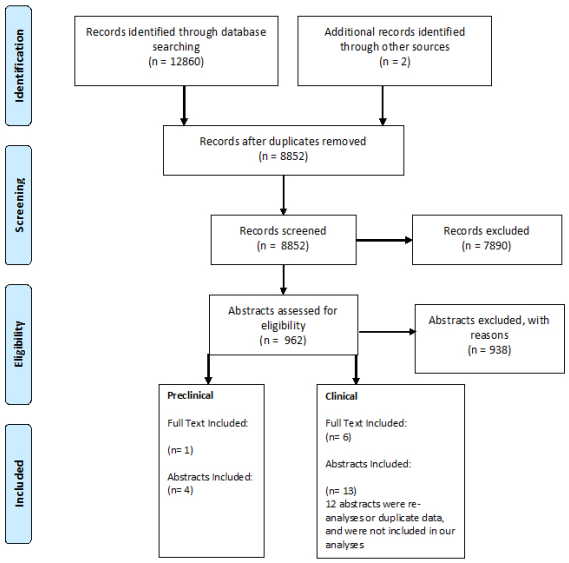

Figure 1. Study selection flow diagram

304x209mm (96 x 96 DPI)

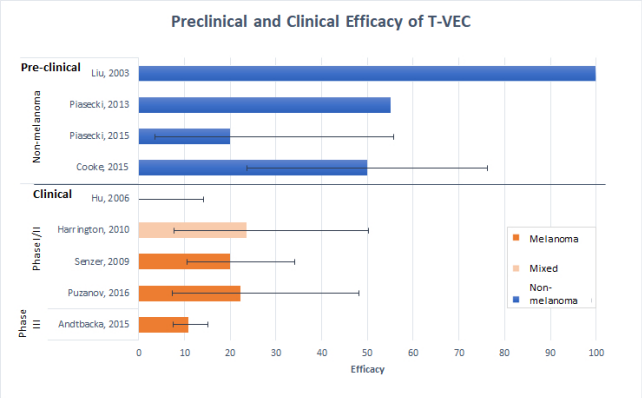

Figure 2. Preclinical and clinical efficacy of T-VEC. Four preclinical studies using mice demonstrated efficacy rates of 20-100% and clinical studies (3 melanoma and 2 mixed malignancy studies) demonstrated efficacy rates from 0-23.5%. Non-melanoma/mixed studies are represented by blue bars whereas melanoma studies are represented by orange bars. Where possible, complete regression rates of contralateral tumors for mice were used in preclinical studies and complete response was used for clinical studies. Efficacy rates decrease in the preclinical to clinical translation and upon more rigorous study design in later phase clinical trials. If possible, error bars are plotted and represent 95% confidence intervals. Some studies were not included in this analysis due to no reporting of outcome for CR.

304x209mm (96 x 96 DPI)

Online Supplementary File 1: Search Strategy

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>
Search Strategy:
---

1 Talimogen* laherparepvec.mp. (28)
2 t vec.mp. (24)
3 OncoVEX*.mp. (15)
4 Imlygic.mp. (2)
5 JS1 34*.tw. (5)
6 or/1-5 (51)
7 Oncolytic Virotherapy/ or Oncolytic Viruses/ or cancer vaccines/tu (7016)
8 (cancer adj2 (vaccine* or virus* or virotherap* or viral therap*)).tw. (5206)
9 exp neoplasms/ or cancer.tw. (3097511)
10 or/7-9 (3097684)
11 simplexvirus/ or herpesvirus 1, human/ or Herpes Simplex/ (32955)
12 (hsv1 or hsv or herpesvirus or Herpes).tw. (72684)
13 11 or 12 (77360)
14 10 and 13 (12929)
15 ((oncolyt* or cancer or tumor or tumour) adj3 (hsv1 or hsv or hsv or herpesvirus or Herpes)).tw. (846)
16 (oncolyt* adj3 (virotherap* or virus* or viral therap*)).tw. (2183)
17 or/14-16 (14671)
18 exp animal experimentation/ or exp models, animal/ or animals/ or mammals/ or vertebrates/ or exp fishes/ or exp amphibia/ or exp reptiles/ or exp birds/ or exp hyraxes/ or exp marsupialia/ or exp monotremata/ or exp scandentia/ or exp chiroptera/ or exp carnivora/ or exp cetacea/ or exp Xenarthra/ or exp elephants/ or exp insectivora/ or exp lagomorpha/ or exp rodentia/ or exp sirenia/ or exp Perissodactyla/ or primates/ or exp strepsirhini/ or haplorhini/ or exp tarsii/ or exp platyrrhini/ or catarrhini/ or exp cercopithecidae/ or gorilla gorilla/ or pan paniscus/ or pan troglodytes/ or exp pongo/ or exp hylobatidae/ or hominidae/ (5893175)
19 (animal$1 or chordata or vertebrate* or fish$2 or amphibian* or amphibium* or reptile$1 or bird$1 or mammal* or dog or dogs or canine$1 or cat or cats or hyrax* or marsupial* or monotrem* or scandentia or bat or bats or carnivor* or cetacea or edentata* or elephant* or insect or insects or insectivore or lagomorph* or rodent$2 or mouse or mice or murine or murinae or muridae or rat or rats or pig or pigs or piglet$1 or swine or rabbit$1 or sheep$1 or goat$1 or horse$1 or equus or cow or cows or cattle or calf or calves or bovine or sirenia or ungulate$1 or primate$1 or prosimian* or haplorhini* or tarsiiform* or simian*or platyrrhini or catarrhini or cercopithecidae or ape or apes or hylobatidae or hominid* or chimpanzee* or gorilla* or orangutan* or monkey or monkeys or ape or apes).tw. (4149175)
20 (preclinic$ or pre clinic$).tw. (72274)
21 or/18-20 (6474971)
22 17 and 21 (6511)
23 6 or 22 (6545)

Online Supplementary File 2. Preclinical Study Characteristics

Author, Year	Year Study Conducted	Country	Study Design	Species	Strain	Model	Type of Cancer	Baseline Tumor Size	Gender	Mean Age	Mean Weight	Co-Interventions	Duration of Follow Up
Cooke, 2015	---	USA	Interventional; Non-Controlled	Mouse	Balb/c	A20 Murine Lymphoma	Lymphoma	150 mm^3	Female	---	---	N/A	---
Piasecki, 2015	---	---	Controlled Comparison	Mouse	C57Bl/6	Syngeneic MC-38 Colon Carcinoma	Colon Cancer	---	---	---	---	Anti-PD-1	---
Piasecki, 2013	---	---	Controlled Comparison	Mouse	---	A20 Syngeneic Contralateral Model	Lymphoma	---	---	---	---	N/A	10 days
Liu, 2003	2002	UK	Controlled Comparison	Mouse	Balb/c	Syngeneic A20 Lymphoma	Lymphoma	0.5 cm diameter	---	---	---	Immunization wild type HSV1	35 days

---: not reported
N/A: not applicable
Cooke, 2016 did not provide any relevant information

Online Supplementary File 3. Preclinical Intervention and Comparator Characteristics

Author, Year	Experiment	Group	N	Dose 1	Frequency Dose 1	Duration Dose 1	Dose 2	Frequency Dose 2	Duration Dose 2	Dose 3	Frequency Dose 3	Duration Dose 3
Cooke, 2015	1	Cohort 1: TVEC	10	$3x10^4$ PFU	---	---	N/A	N/A	N/A	N/A	N/A	N/A
		Cohort 2:	10	---	---	---	N/A	N/A	N/A	N/A	N/A	N/A
		Cohort 3: TVEC	10	$3x10^6$ PFU	Every 3 days	1 week	N/A	N/A	N/A	N/A	N/A	N/A
		Cohort 4:	10	---	---	---	N/A	N/A	N/A	N/A	N/A	N/A
Piasecki, 2015	1	Int: OncoVEXmuGM-CSF + Anti-PD1	---	T-VEC: ---	Every 3 days	3 doses	Anti-PD-1: ---	Twice per wk	---	N/A	N/A	N/A
		Int: OncoVEXmuGM-CSF	---	T-VEC: ---	---	---	---	---	---	---	---	---
		Con: Anti-PD-1	---	Anti-PD-1: ---	---	---	---	---	---	---	---	---
Piasecki, 2013	1	Int: T-VEC	---	$5x10^6$ PFU	single	single	N/A	N/A	N/A	N/A	N/A	N/A
		Con: Vehicle	---	---	single	single	N/A	N/A	N/A	N/A	N/A	N/A
Liu, 2003	1	Int: JS1/34.5-/47-/mGM-CSF	10	10^6 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
			10	10^7 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
			10	10^8 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
		Int: JS1/34.5-/47-	10	10^6 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
			10	10^7 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
			10	10^8 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
		Con: vehicle	10	50µl	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
	2	Int: JS1/34.5-/47-/mGM-CSF	10	10^8 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
		Con: Vehicle	10	50µl	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A

All doses of T-VEC were given by injection intratumorally
---: not reported
Int: intervention; Con: control; wk: week; PFU: plaque forming units
Cooke, 2016 did not provide any relevant information

Online Supplementary File 4.. Clinical Study Characteristics

Author, Year	Country	Year Study Conducted	Study Type	Type of Cancer	Primary Outcomes	Secondary Outcomes
Andtbacka, 2015	USA, UK, Canada and South Africa	2009-2014	Interventional; Randomized (OPTiM Trial)	Melanoma	Efficacy: DRR	Efficacy: ORR OS Best Overall Response Onset and duration of response Time to treatment failure
Long 2015	USA, Australia, Switzerland, Spain	2014-2022	Interventional: Non-randomized, No control	Melanoma	Safety: Dose Limiting Toxicities	Efficacy: DRR OS Progression Free Survival Safety: AEs
Puzanov, 2015	USA	2013-2014	Interventional: Non-Randomized	Melanoma	Safety: Dose Limiting Toxicities	Efficacy: ORR Safety: Grade ≥3 AEs
Chang, 2012	USA	2006-2008	Interventional: Non-randomized, No control	Pancreatic Cancer	Efficacy: Detection of T-VEC in blood and urine Presence of Anti-HSV1 Antibodies Safety: AEs	Efficacy: ORR Change in sum of longest tumor diameter Change in pain intensity
Harrington, 2010	UK	2005-2010	Interventional: Non-randomized, No control	Squamous Cell Carcinoma	Safety: AEs	Efficacy: Antitumor Activity OS* Complete Response* Partial Response* Progression Free Survival*
Senzer 2009	USA	2005-2008	Interventional; (non-controlled, non randomized)	Melanoma	Efficacy: ORR	Efficacy: OS Safety: AEs
Hu, 2006	USA, UK	---	Interventional: Non-randomized, No control	Breast, Colorectal, Melanoma, Head and Neck	Efficacy: Biodistribution Safety: AEs	Efficacy: GM-CSF expression HSV antigen associated necrosis Viral Replication Local Reactions

---: Not Reported
*: not reported a priori
AEs - adverse events; CHN– cutaneous head and neck; DRR – durable response rate; ECOG – Eastern Cooperative Oncology Group; ORR- objective response rate; OS – overall survival

Online Supplementary File 5. Clinical Patient Characteristics

Author, Year	Group	Patients (N)	Median Age (range)	Sex (n; F)	Metastasis Stage (n; Stage IVM1b/c)	Line of Therapy (n; first line)	HSV Serostatus (n; Seropositive, n; unknown)
Andtbacka, 2015	T-VEC	295	63 (22-94)	122	131	138	175, 23
Long, 2015	T-VEC + Pembrolizumab	21	58	13	11	---	---
Puzanov, 2015	T-VEC + Ipilimumab	18	---	---	---	18	---
Chang, 2012	T-VEC	17	54	6	---	---	---
Harrington, 2010	T-VEC and Chemo radiotherapy	17	58 (41-74)	2	3	---	---
Senzer, 2009	T-VEC	50	62 (34-88)	28	24	0	36, 1
Hu, 2006	T-VEC	30	55 (30-80)	23	---	0	19

--- : Not Reported

Online Supplementary File 6. Clinical Intervention and Comparator Characteristics

Author, Year	Arm	Dose 1	Time of Dose 1	Frequency of Dose 1	Dose 2	Time of Dose 2	Frequency of Dose 2	Dose 3	Time of Dose 3	Frequency of Dose 3	Intervention Window	Follow Up Duration
Andtbacka, 2015	T-VEC	10^6 PFU/ml (≤4ml)	Week 1	single	10^8 PFU/ml (≤4ml)	Week 4	Q2W	N/A	N/A	N/A	Median: 23 wks (0.1-79 wks)	Median: 44 mo (32-58 mo)
	GM-CSF	125 µg/m^2	Week 1	Once daily 14/28 day cycles	N/A	N/A	N/A	N/A	N/A	N/A	Median: 10 wks (0.6 to 72 wks)	---
Long, 2015	T-VEC + Pemb.	TVEC: 10^6 PFU/ml	Day 1	single	T-VEC: 10^8 PFU/ml	Day 22	Q2W	Pemb: 200 mg	Day 36	Q2W	Median TVEC: 13 wks Median Pemb: 10 wks	---
Puzanov, 2015	T-VEC + Ipilimumab	TVEC: 10^6 PFU/ml (≤4ml)	Week 1	single	TVEC: 10^8 PFU/ml (≤4ml)	Week 4	Q2W	Ipilimumab: 3mg/kg	Week 6	Q3W	TVEC: until DLT Ipi: 12 wks	17 mo minimum
Chang, 2012	Cohort 1	10^4 PFU/ml	Week 1*	single	10^5 PFU/ml	Week 4*	Q3W*	N/A	N/A	N/A	up to 15 wks	---
	Cohort 2	10^5 PFU/ml	Week 1	single	10^6 PFU/ml	Week 4	Q3W	N/A	N/A	N/A	up to 15 wks	---
	Cohort 3	10^6 PFU/ml	Week 1	single	10^7 PFU/ml	Week 4	Q3W	N/A	N/A	N/A	up to 15 wks	---
Harrington, 2010	Cohort 1	T-VEC: 10^6 PFU/ml	Day 1	Q3W	Cisplatin: 100mg/m^2	Day 1	Q3W	N/A	N/A	N/A	Up to 9 weeks	Median: 29mo (19-40mo)
	Cohort 2	T-VEC: 10^6 PFU/ml	Day 1	single	T-VEC: 10^7 PFU/ml	Day 22	Q3W	Cisplatin: 100mg/m^2	Day 1	Q3W	Up to 9 weeks	Median: 29mo (19-40mo)
	Cohort 3	T-VEC: 10^6 PFU/ml	Day 1	single	T-VEC: 10^8 PFU/ml	Day 22	Q3W	Cisplatin: 100mg/m^2	Day 1	Q3W	Up to 9 weeks	Median: 29mo (19-40mo)
	Cohort 4	T-VEC: 10^6 PFU/ml	Day 1	single	T-VEC: 10^8 PFU/ml	Day 22	Q3W	Cisplatin: 100mg/m^2	Day 1	Q3W	Up to 9 weeks	Median: 29mo (19-40mo)
Senzer, 2009	T-VEC	10^6 PFU/ml (≤4ml)	Week 1	Single	10^8 PFU/ml (≤4ml)	Week 4	Q2W	N/A	N/A	N/A	Max: 48 wks Median: 11 wks	Median: 18 mo (11-36 mo)
Hu, 2006	Single Dose Group 1	10^6 PFU/ml	---	single	N/A	N/A	N/A	N/A	N/A	N/A	Single dose	6 wks
	Single Dose Group 2	10^7 PFU/ml	---	single	N/A	N/A	N/A	N/A	N/A	N/A	Single dose	6 wks
	Single Dose Group 3	10^8 PFU/ml	---	single	N/A	N/A	N/A	N/A	N/A	N/A	Single dose	6 wks
	Multi-dose Group 1	10^6 PFU/ml	---	single	10^7 PFU/ml	---	Q1-3W	N/A	N/A	N/A	3-9 wks*	6 wks post final injection
	Multi-dose Group 2	10^6 PFU/ml	---	single	10^8 PFU/ml	---	Q1-3W	N/A	N/A	N/A	3-9 wks*	6 wks post final injection

Author, Year	Arm	Dose 1	Time of Dose 1	Frequency of Dose 1	Dose 2	Time of Dose 2	Frequency of Dose 2	Dose 3	Time of Dose 3	Frequency of Dose 3	Intervention Window	Follow Up Duration
	Multi-dose Group 3	10^8 PFU/ml	---	Q1-3W	N/A	N/A	N/A	N/A	N/A	N/A	3-9 wks*	6wks post final injection

DLT: Dose Limiting Toxicity; Pemb: Pembrolizumab; Q2W: every two weeks Q3W: every three weeks; Q1-3W: every 1-3 weeks; Q6W every 6 weeks
--- : not reported
T-VEC was given by intra-tumoral injection in all studies

Online Supplementary File 7. Preclinical Efficacy Data

Author, Year	Experiment	Group	N – Animals Studied	N – Lesions Studied	Baseline Mean Tumor Measure (Standard Error of Mean)	Final Mean Tumor Measure (Standard Error of Mean)	CR - Injected	CR - Contralateral	Duration of Follow Up
Cooke, 2015	1	INT: TVEC $3x10^4$ PFU	10	---	---	---	---	---	---
		INT: TVEC $3x10^6$ PFU	10	10	~150mm^3	---	10/10	5/10	---
		---	10	---	---	---	---	---	---
		---	10	---	---	---	---	---	---
Piasecki, 2015	1	INT: OncoVexmuGM-CSF	---	---	---	---	---	---	---
		INT: OncoVexmuGM-CSF + Anti-PD-1	---	20	---	---	8/10	2/10	---
		CON: Anti Pd-1	---	---	---	---	---	---	---
Piasecki, 2013	1	INT: T-VEC	---	---	---	---	70-100%	50-60%	10 days
		CON: Vehicle	---	---	---	---	---	---	---
Liu, 2003	1	INT: JS1/34.5-/47-/mGM-CSF; injected	10	N/A	5.2mm (0.34)	0.004mm (0.31)	N/A	N/A	22 days
		INT: JS1/34.5-/47-/mGM-CSF; uninjected	10 (same as injected)	N/A	5.7mm (0.29)	1.1 mm (0.73)	N/A	N/A	22 days
		INT: JS1/34.5-/47-; injected	10	N/A	5.4mm (0.37)	1.4 mm (1.36)	N/A	N/A	22 days
		INT: JS1/34.5-/47-; uninjected	10 (same as injected)	N/A	6.2mm (0.29)	5.4mm (2.01)	N/A	N/A	22 days
		CON: Vehicle; injected	10	N/A	5.4mm (0.40)	11.9mm (2.69)	N/A	N/A	22 days
		CON: Vehicle; uninjected	10 (same as injected)	N/A	5.6mm (0.46)	13.2mm (2.76)	N/A	N/A	22 days
	2	INT: JS1/34.5-/47-/mGM-CSF	10	N/A	5.5mm (0.34)	2.2mm (1.6)	N/A	N/A	21 days
		CON: Vehicle	10	N/A	5.6mm (0.23)	13.8mm (1.2)	N/A	N/A	21 days

--- : not reported
INT: intervention
CON: control
Liu 2003 data from experiment 1 taken from 10^8 dose
Cooke, 2016 did not provide any relevant information

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	4
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5-6
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	5-6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	7

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	6
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	7-8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9-10
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	8-9
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	8-9
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	9-10
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11-12
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	12-13
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	13
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	14

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

BMJ Open

Mapping the preclinical to clinical evidence and development trajectory of the oncolytic virus talimogene laherparepvec (T-VEC): a systematic review

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-029475.R2
Article Type:	Original research
Date Submitted by the Author:	01-Nov-2019
Complete List of Authors:	Lalu, Manoj; Ottawa Hospital Research Institute, Clinical Epidemiology Leung, Garvin; Ottawa Hospital Research Institute, Clinical Epidemiology Dong, Yuan Yi; Ottawa Hospital Research Institute, Clinical Epidemiology Montroy, Joshua; Ottawa Hospital Research Institute, Clinical Epidemiology Butler, Claire; Ottawa Hospital Research Institute, Clinical Epidemiology Auer, Rebecca; Ottawa Hospital Research Institute, Clinical Epidemiology Fergusson, Dean; Ottawa Hospital Research Institute, Medicine
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Oncology
Keywords:	TVEC, oncolytic virus, cancer, translation, review

Mapping the preclinical to clinical evidence and development trajectory of the oncolytic virus talimogene laherparepvec (T-VEC): a systematic review

Manoj M. Lalu† MD, PhD, FRCPC[1,2,3,4], Garvin J. Leung† MPH[2,5], Yuan Yi Dong[2,5], Joshua Montroy MSc,[2] Claire Butler[2], Rebecca C. Auer MD, MSc, FRCSC[7,8], Dean A. Fergusson PhD, MHA*[2,6,7]

[1]Department of Anesthesiology and Pain Medicine, The Ottawa Hospital, University of Ottawa, Ottawa, ON, Canada
[2]Blueprint Translational Research Group, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
[3] Regenerative Medicine Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
[4] Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada
[5]Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada
[6]School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada
[7]Department of Surgery, The Ottawa Hospital, University of Ottawa, Ottawa, ON, Canada
[8]Cancer Therapeutics Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
† These authors contributed equally to this work

Emails: MML: mlalu@toh.ca; GJL: garvin.leung@uottawa.ca; YYD: ydong044@uottawa.ca; JM: jmontroy@ohri.ca; CB: clbutler@toh.ca; RCA: rauer@toh.ca; DAF: dafergusson@ohri.ca

*Corresponding author:
Dean A. Fergusson email: dafergusson@ohri.ca
Centre for Practice-Changing Research,
Office L1298a
501 Smyth Road, Box 201B
Ottawa, Ontario, Canada K1H 8L6
Tel. 1-613-737-8480
Fax. 1-613-739-6938

Running Head: The efficacy of T-VEC: A systematic review

Word count: 3276

ABSTRACT

Objective: This study aimed to conduct a systematic review of preclinical and clinical evidence to chart the successful trajectory of Talimogene laherparepvec (T-VEC), from the bench to the clinic.

Design: This study was a systematic review. The primary outcome of interest was the efficacy of treatment, determined by complete response. Abstract and full-text selection as well as data extraction was done by two independent reviewers. The Cochrane risk of bias tool was used to assess the risk of bias in studies.

Setting: Embase, Embase Classic, and OvidMedline were searched from inception until May 2016 to assess its development trajectory to approval in 2015.

Participants: Preclinical and clinical controlled comparison studies, as well as observational studies.

Interventions: T-VEC for treatment of any malignancy.

Results: 8,852 records were screened and five preclinical (n=150 animals) and seven clinical studies (n=589 patients) were included. We saw large decreases in T-VEC's efficacy as studies moved from the laboratory to patients, and as studies became more methodologically rigorous. Preclinical studies reported complete regression rates up to 100% for injected tumors and 80% for contralateral tumors, while the highest degree of efficacy seen in the clinical setting was a 24% complete response rate, with one study experiencing a complete response rate of 0%. We were unable to reliably assess safety due to the lack of reporting, as well as the heterogeneity seen in adverse event definitions. All preclinical studies had high or unclear risk of bias, and all clinical studies were at a high risk of bias in at least one domain.

Conclusions: Our findings illustrate that even successful biotherapeutics may not demonstrate a clear translational road map. This emphasizes the need to consider increasing rigour and transparency along the translational pathway.

PROSPERO Registration Number: CRD42016043541

Keywords: T-VEC, oncolytic virus, cancer, translation, review

Strengths and limitations of this study

- Comprehensive, up-to-date review of the efficacy and safety of T-VEC
- Threats to both internal validity and construct validity were performed
- Reporting of methods and findings was incomplete in most of the studies included
- Poor reporting and study design are major contributors to the ongoing reproducibility crisis in preclinical research

BACKGROUND

Preclinical research receives approximately half of the world's biomedical research funding, yet very few of its findings translate clinically. This represents an enormous waste of resources with an estimated 28 billion dollars per year in the US alone being spent on biomedical research which is not reproducible and therefore not translatable.(1) One study found that only 5% of highly efficacious preclinical therapeutics were clinically translated.(2) These successes often take almost twenty years to become successfully translated across the research spectrum. (2, 3)

Although the process of clinical translation is complicated, the transition from bench-to-bedside often starts with preclinical research. These investigations (usually on animals or cells), are aimed at studying efficacy, pharmacokinetics and dynamics, as well as detailing safety.(4) Next, a drug is tested in a phase I clinical trial, which usually contains a small number of participants and is aimed at studying the safety of the drug. If a drug is safe, it may proceed to phase II which are larger than phase I studies and are designed to test safety, pharmacokinetics, pharmacodynamics, and optimal dosing regimens. They may also offer preliminary evidence of drug efficacy. Finally, a methodologically rigorous phase III study is performed. These studies are designed and powered to test efficacy in the patient population of interest (usually against a comparator such as placebo), as well as identify rarer adverse events which may have gone unnoticed in a smaller phase I or II study.(5)

Given the high failure rate in translating therapies across this spectrum, as well as significant time-lags associated with translation, it is important that we examine the few agents that have successfully crossed the preclinical-to-clinical bridge in order to learn from and replicate their success. Thus, we conducted a comprehensive evaluation of available evidence supporting the successful translation of Talimogene laherparepvec (T-VEC). T-VEC is a modified HSV-1 virus

produced by *Amgen* and it is the first, and only, FDA approved oncolytic virus therapy; it is currently approved to treat advanced melanoma.(6) Oncolytic viruses are an emerging cancer therapy that work by preferentially targeting and infecting cancer cells.(6) Upon infection, oncolytic viruses can induce an anti-tumor immune response that reduces tumor burden. TVEC was chosen as a model due to the fact that it is the only approved oncolytic virus therapy to date, despite the multitude of agents under investigation.(7)

Through a careful evaluation of T-VEC development we hoped to identify factors that may contribute to bench-to-bedside success. This may serve an exemplar for other therapies as they move along the translational continuum. Thus, the purpose of this systematic review was to map the successful preclinical to clinical trajectory of T-VEC to inform the development paths of future biotherapeutics.

METHODS

Our review was registered in full on PROSPERO, the international prospective register of systematic reviews (no. CRD42016043541). The review is reported in accordance to the PRISMA guidelines.(8)

Eligibility Criteria

We included all clinical and preclinical *in vivo* controlled comparison studies of T-VEC for treatment of any malignancy (randomized, pseudo-randomized, and non-randomized studies), as well as observational studies such as case-control, case-series and case reports. Studies reporting only *ex vivo* or *in vitro* experiments were excluded. For both preclinical and clinical studies, we included studies that administered T-VEC as a monotherapy or in combination with other therapies for treatment of malignancy. We had no exclusions on comparison treatments, which include standard line therapy or no treatment.

Outcomes

The primary outcome of interest was the efficacy of treatment. Our primary indicator of efficacy was complete response. Other measures of efficacy such as survival, response rates (durable, partial, objective), time to treatment failure, and disease stability were also collected. Such measures were based on the Response Evaluation Criteria in Solid Tumors (RECIST) Guidelines.(9) In preclinical studies, additional measures of efficacy such as changes in mean tumor volume and number of lesions were collected. The primary indicator of efficacy, complete response, was used as the primary outcome regardless of reporting within the individual study, in order to assess the continuity of evidence along the research spectrum. The secondary outcome of

interest was safety, for which we collected data on all adverse events in preclinical and clinical studies.

Literature Search

In collaboration with a medical information specialist (Risa Shorr, Learning Services, The Ottawa Hospital) a search strategy was designed to identify all relevant preclinical and clinical studies. Searches were conducted in the following databases: Embase, Embase Classic, and OvidMedline from inception until May 2016. This time frame was chosen to ensure all published studies that contributed T-VECs FDA approval in 2015 were included. Search terms included: Talimogen laherparepvec, Tvec, OncoVEX and Imlygic. Additional terms pertaining to preclinical studies (e.g. animal experiment/model) and oncology (e.g. cancer, neoplasm, oncolytic virus) were also included. Studies were also screened for inclusion based on reference tracking, by scanning the bibliography of included primary studies and relevant review articles. We did not impose any restrictions on language or publication type. A grey literature search was not performed. The finalized search strategy can be found in online supplementary file 1.

Study Selection Process

Studies identified by our literature search were collated and duplicates were removed. Titles and abstracts were independently screened for inclusion by two reviewers using DistillerSR (Evidence Partners, Ottawa, ON). Those deemed potentially relevant were recorded, and full-text articles were obtained. The same reviewers screened full articles for final eligibility. Disagreements at any stage were resolved by discussion or by consultation with a senior team member when necessary. The study selection process was documented using a PRISMA flow diagram (Figure 1).

Data Extraction

All data extraction was completed independently and in duplicate, using a standardized and piloted data extraction form, with disagreements resolved as mentioned above. Data pertaining to general and intervention characteristics of the included studies were extracted (e.g. study design, country, type of malignancy, dosing of intervention and comparator treatments). For clinical studies, data was collected on patient characteristics (e.g. age, sex, cancer staging, HSV status). For preclinical studies, characteristics on the animal model were extracted (e.g. type of species, cell line used, disease induction method, age, sex, weight).

Risk of bias – assessment to risk of internal validity

Clinical studies that met inclusion criteria were assessed for risk of bias in duplicate, according to the recommended methodology of the Cochrane Collaboration.(10) Five types of biases (selection, performance, detection, attrition, reporting biases) were assessed using six domains: randomization, allocation concealment, blinding of participants/personnel, outcome assessment blinding, incomplete outcome reporting, and selective outcome reporting. Additional domains assessed for risk of bias were: i) reported conflicts of interest, ii.) sample size calculation, and iii.) funding. Each domain was given a score of "high", "unclear", or "low" risk of bias for each study. Risk of bias assessment for preclinical studies were assessed using a modified Cochrane Risk of Bias tool and assessed the same domains as indicated for clinical studies.(11)

Assessment of threats to construct validity

Construct validity is the concept in how much a preclinical experiment (i.e. animal studies) corresponds to the clinical entity it is intended to model. There are various threats to construct validity that can be introduced from the preclinical study design. The items evaluated in duplicate

for each preclinical study include: i.) use of adult animals, ii.) use of animals with advanced stage disease (defined as the presence of multiple visceral lesions and/or clinical/histological signs of malignant progression), iii.) immune status of animals to HSV, iv.) whether a xenograft model was used, and v.) the use of a humanized immune system model. Each of these items was given a score of "yes", "no", or "unclear" for every preclinical study.

Statistical Analysis

Efficacy was expressed as proportions with accompanying 95% confidence intervals. If confidence intervals were not present within the individual study, they were calculated via standard methods.(12) To assess the continuity between preclinical and clinical studies, the efficacy of studies was plotted as percentage response.

Deviations from Protocol

We were unable to assess safety as we could not acquire patient-level safety data. Furthermore, our primary efficacy outcome stated in protocol was durable response rate. However, this was changed to complete response as most clinical studies did not report durable response and we needed to track TVEC's trajectory over several studies. We acknowledge the limitation of this approach, given the FDA approved TVEC based on the OPTIM trial,(13) the primary endpoint of which was durable response rate. Subgroup analyses, meta-analyses, Egger's test, and pooling of data could not be conducted due to the limited available data.

Patient and Public Involvement

Patients and the public were not involved in this research.

RESULTS

Upon removal of duplicates, a total of 8,852 references were identified by the electronic search. During the review of titles and abstracts, 7,890 references were excluded. Following full text screening, another 938 articles were excluded for reasons such as wrong study design (i.e. review article), or wrong study intervention (i.e. a different cancer therapeutic). A total of 7 clinical studies,(13-19) and 5 preclinical studies(20-24) were included in our review (Figure 1).

Characteristics of Included Trials

Characteristics of included studies are shown in Table 1. Preclinical studies were published between 2003 and 2016 and sample sizes ranged from 20 to 90. Of the 5 preclinical studies, three used a lymphoma model, one used a colorectal model, and one used a melanoma model. All studies were performed in mice. The duration of follow-up was reported by two studies and ranged from 10 days to 35 days. The dose of T-VEC used ranged from $3x10^4$ plaque forming units (PFU) to $5x10^6$ PFU. One frequency of T-VEC administration varied from, every three days for one week, every three days for nine days, a single dose given only once, and every other day for five days. Specific details of study and intervention characteristics for each preclinical study can be found in online supplementary files 2 and 3.

Clinical studies were published between 2006 and 2016 and took place in seven countries. Sample sizes ranged from 17 to 295. Of the seven clinical studies, four were in melanoma patients, one was in pancreatic cancer patients, one in head and neck cancer patients and one studied a mix of breast, colorectal, melanoma and head and neck cancer patients. Six were either Phase I or II, and one trial was a Phase III evaluation. The primary outcome was efficacy in two studies, safety in three studies and a combination of efficacy and safety in the other two studies. The duration of follow-up ranged from six weeks to 44 months.

T-VEC was administered alone in four studies, while it was administered immediately following to systemic therapy in 3 studies. The dose of T-VEC administered ranged from 10^4 PFU/mL to 10^8 PFU/mL. In the large, Phase III study, T-VEC was administered at ≤4mL x 10^6 PFU/mL once, and then three weeks later, ≤4mL 10^8 PFU/mL was administered every two weeks for a median of 23 weeks. A similar dosing regimen was used in three other trials. The other three trials were dose-finding in nature and had multiple trial arms receiving increasing doses of T-VEC. In-depth study details, as well as participant and intervention details for each study can be found in online supplementary files 4, 5, and 6.

Efficacy of Treatment

Treatment efficacy for each study is summarized in Table 1 and Figure 2. Preclinical studies reported complete regression rates up to 100% for injected tumors and 80% for contralateral tumors (see also online supplementary file 7). In comparison, the first published Phase I T-VEC clinical trial reported a complete response of 0% for cutaneous lesions caused by malignancies of head and neck, breast, colorectal, and melanoma. Of the multiple malignancies treated, melanoma had the best response in this trial. Subsequent Phase I/II melanoma trials were then conducted and demonstrated complete response rates of 20-22%. This was followed by the Phase III OPTIM melanoma trial, which had a complete response rate of 10.8%. Studies involving non-melanoma cancers varied with efficacies between 0-24%

Safety of treatment

We attempted to assess safety across clinical studies, however we were unable to obtain patient level data from any of the studies. The definitions of adverse events, and the manner in which they were classified, was found to be highly heterogenous across studies. Studies did not specify what

percent of adverse events were repeated adverse events from the same patient(s), used different criteria for recording and reporting adverse events, categorized them differently, etc. Therefore, we were unable to pool adverse events or interpret findings reliably.

Validity Assessments

Construct validity, the concept of how well an animal model represents the clinical entity it is intended to mimic, was first assessed through the following domains: the use of appropriately-aged mice, advanced stage of disease, HSV-immunity, and types of mouse models. None of the preclinical studies fully reported or used methodologies to reduce threats to construct validity domains (Table 2). No studies declared using adult animal models, no studies used animals with late stage disease, only one study used animals immune to HSV, no studies used a xenograft model, and no studies reported using an animal model with a humanized immune system.

We also assessed internal validity (i.e. risk of bias) and found that all preclinical studies had high or unclear risk of bias across the assessed domains: randomization sequence, allocation concealment, blinding, incomplete reporting, sample size calculation, and funding source (Table 3). For clinical studies, early phase trials had high or unclear risk of bias across at least six of nine domains whereas the more robust Phase III OPTIM trial had the lowest risk of bias and also the lowest efficacy of any of the published melanoma clinical trials (Table 4). Reporting of key methodological elements was lacking.

DISCUSSION

We hoped to synthesize the evidence to produce a clear road map of T-VEC's translation in the published literature to follow the journey of a successful biotherapeutic, to be used as a blueprint for similar efforts in the future. Yet, we were unable to paint a clear picture of how the evidence

was utilized in proceeding to melanoma clinical trials. Rather, our assessment uncovered a disconnect between *in vivo* preclinical and clinical findings. Furthermore, the road map was plagued with poor reporting, high risk of bias, and insufficient data along the translational path. Overall, we were surprised by the pace and magnitude of diminishing efficacy as T-VEC moved from bench to bedside and then towards later phase clinical trials (i.e. Phase I to III). Although T-VEC was successful in terms of gaining regulatory approval, its translational path is complicated, and the pieces of the evidence puzzle do not easily fit together. While we appreciate that translation is not a predictable linear process, it is difficult to learn from the example of T-VEC given the available and reported pre-clinical and clinical evidence.

While many novel therapeutics are under intellectual property rights, details of study design and results should be transparently reported for scientists, clinicians, and patients to evaluate findings. The fact that the only FDA approved oncolytic virus therapy is not clearly reported illustrates the issues plaguing the success of cancer therapeutics. Nonetheless, T-VEC has shown some efficacy in treating refractory melanoma(25) and numerous clinical trials are underway to assess its use in combination with other cancer regimens and in treating other malignancies. It is also the recommended treatment by the National Comprehensive Cancer Center for patients with in-transit melanoma.(26)

Perhaps the largest discrepancy noted was that only a single preclinical study used a melanoma model, whereas all but two clinical studies administered T-VEC to melanoma patients. Conversely, lymphoma, which was used in three preclinical studies, was not assessed in clinical studies. Interestingly, our subsequent searches found that Amgen's FDA filing (STN# 125518.000) (27) for T-VEC did not appear to report on any *in vivo* melanoma models, whereas the EMA report did (EMA/734400/2015).(28) Thus, the majority of animal models were off-target from the

malignancies studied in clinical trials and may have poorly represented melanoma in the clinical setting. Coupled with these findings was the fact that the majority of our studies were found to be at a high risk of bias. Such threats to internal validity can bias results and may help explain T-VEC's superior preclinical efficacy compared to later phase clinical trials. A lack of randomization and blinding in preclinical studies has been associated with inflated effect sizes,(29, 30) thus this may partially explain the preclinical to clinical discrepancy of T-VEC.

Reporting of methods and findings was incomplete in most of the studies included. Only one full preclinical article on T-VEC was published, and solely aggregate patient data for later phase trials was available. Poor reporting and study design are major contributors to the ongoing reproducibility crisis in preclinical research.(31) Thus, in hopes of presenting a clearer picture of T-VEC's successful translation, we contacted *Amgen* to obtain preclinical *in vivo* melanoma data, patient-level safety data, and any additional efficacy data. Patient-level data would afford the ability to combine data across T-VEC's clinical development and also provide clarification into the categorization of adverse events. Recently, release of individual patient data to third parties has been advocated by the Institute of Medicine, journal editors, and others as it enhances transparency, enables re-analyses of data, and helps address reproducibility.(32) The reporting of harms in clinical trials remains an issue in the scientific community,(33-35) and represents a roadblock to translational success. Some basic steps required to improve the reporting of safety in translational research include the development of standardized scales and instruments, instituting active rather than passive surveillance for toxicity, including detailed information on participant withdrawals due to toxicity, reporting the timing, frequency, and duration of clinically relevant events, and the publication of raw data.(36, 37) *Amgen,* however, was unwilling to enter a data sharing agreement, as they stated that there was little value to compel a transparent data release

for our proposed analyses. This lack of transparency and incomplete reporting is disappointing, especially considering that it was *Amgen* that previously highlighted poor reporting as contributing to its own failure to reproduce 47 of 53 high-impact preclinical cancer studies.(38) Their findings fuelled a call by the NIH and other stakeholders to enhance the reproducibility and transparency of preclinical research.(39)

As stated, we recognize that translation is not a linear process, but we should observe consistent and coherent patterns. Moving forward, we suggest that preclinical and clinical studies for emerging therapies should be fully reported and attention should be given to validities in order to develop more precise estimates of effect early in development. Investigators should carefully match their preclinical model to the intended clinical population; when possible, both disease states and outcomes measured should have high construct validity. Following successful exploratory preclinical studies, investigators should consider preclinical systematic reviews(40) and designing methodologically rigorous confirmatory and/or multicenter preclinical studies.(41) These steps may allow preclinical testing to more accurately forecast downstream clinical results in human patients.(30) Within the trajectory of clinical development (i.e. once clinical trials have been initiated), careful consideration of methods to reduce bias should also be considered (although, this may not be possible for the earliest phase trials). We believe these steps will provide unbiased and valuable information that will ultimately provide patients with cancer therapies that match their preclinical and early clinical promise.

CONCLUSIONS

The findings from our systematic review demonstrate that even successful biotherapeutics may not demonstrate a clear translational road map. The magnitude of efficacy of T-VEC demonstrated in preclinical studies was considerably larger when T-VEC was moved to the clinic, and the most methodologically rigorous trial included in our review demonstrated the smallest degree of efficacy. Methodologically rigorous studies should be performed earlier on in the translational pathway, which may help to get a realistic estimate of treatment efficacy prior to clinical translation.

DECLARATIONS

Ethics approval: Not applicable.

Consent for publication: Not applicable.

Availability of data: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding: Biotherapeutics for Cancer Treatment (BioCanRx) supported the conduct of this study by a Catalyst grant. BioCanRx is a Government of Canada funded Networks of Centres of Excellence and was not involved in any other aspect of the project, such as the design of the project's protocol and analysis plan, the collection of data and analyses. CB was also supported by a BioCanRx studentship. ML is supported by The Ottawa Hospital Anesthesia Alternate Funds Association and the Scholarship Protected Time Program, Department of Anesthesiology and Pain Medicine, uOttawa.

Author contributions: MML and DAF conceptualized the study. MML, RA, DAF, and GJL contributed to the study design. GJL, YYD, JM, and CB conducted data extraction. All authors analysed and interpreted the data. MML and GJL were responsible for drafting the manuscript.

All authors critically reviewed the manuscript and provided intellectual content. All authors approve the final version of the manuscript.

Acknowledgments: We thank Risa Shorr (Ottawa Hospital Research Institute librarian) for systematic search assistance. MML is supported by The Ottawa Hospital Anesthesia Alternate Funds Association and the Scholarship Protected Time Program, Department of Anesthesiology and Pain Medicine, uOttawa.

REFERENCES

1. Freedman LP, Cockburn IM, Simcoe TS. The Economics of Reproducibility in Preclinical Research. PLoS Biol. 2015;13(6):e1002165.
2. Contopoulos-Ioannidis DG, Ntzani E, Ioannidis JP. Translation of highly promising basic science research into clinical applications. Am J Med. 2003;114(6):477-84.
3. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. J R Soc Med. 2011;104(12):510-20.
4. Use of Laboratory Animals in Biomedical and Behavioral Research. Washington (DC)1988.
5. Umscheid CA, Margolis DJ, Grossman CE. Key concepts of clinical trials: a narrative review. Postgrad Med. 2011;123(5):194-204.
6. Rehman H, Silk AW, Kane MP, Kaufman HL. Into the clinic: Talimogene laherparepvec (T-VEC), a first-in-class intratumoral oncolytic viral therapy. J Immunother Cancer. 2016;4:53.
7. Fergusson DA, Wesch NL, Leung GJ, MacNeil JL, Conic I, Presseau J, et al. Assessing the Completeness of Reporting in Preclinical Oncolytic Virus Therapy Studies. Mol Ther Oncolytics. 2019;14:179-87.
8. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009;62(10):1006-12.
9. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47.
10. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.
11. Hooijmans CR, Rovers MM, de Vries RBM, Leenaars M, Ritskes-Hoitinga M, Langendam MW. SYRCLE's risk of bias tool for animal studies. BMC Medical Research Methodology. 2014;14(1):43.
12. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Statistics in medicine. 1998;17(8):857-72.
13. Andtbacka RH, Kaufman HL, Collichio F, Amatruda T, Senzer N, Chesney J, et al. Talimogene Laherparepvec Improves Durable Response Rate in Patients With Advanced Melanoma. J Clin Oncol. 2015;33(25):2780-8.
14. Hu JC, Coffin RS, Davis CJ, Graham NJ, Groves N, Guest PJ, et al. A phase I study of OncoVEXGM-CSF, a second-generation oncolytic herpes simplex virus expressing granulocyte macrophage colony-stimulating factor. Clinical cancer research : an official journal of the American Association for Cancer Research. 2006;12(22):6737-47.
15. Senzer NN, Kaufman HL, Amatruda T, Nemunaitis M, Reid T, Daniels G, et al. Phase II clinical trial of a granulocyte-macrophage colony-stimulating factor-encoding, second-

generation oncolytic herpesvirus in patients with unresectable metastatic melanoma. J Clin Oncol. 2009;27(34):5763-71.

16. Long GV, Dummer R, Ribas A, Puzanov I, Michielin O, Vanderwalde A, et al. 24LBA Safety data from the phase 1b part of the MASTERKEY-265 study combining talimogene laherparepvec (T-VEC) and pembrolizumab for unresectable stage IIIB-IV melanoma. European Journal of Cancer. 2015;51:S722.
17. Puzanov I, Milhem MM, Minor D, Hamid O, Li A, Chen L, et al. Talimogene Laherparepvec in Combination With Ipilimumab in Previously Untreated, Unresectable Stage IIIB-IV Melanoma. J Clin Oncol. 2016;34(22):2619-26.
18. Harrington KJ, Hingorani M, Tanay MA, Hickey J, Bhide SA, Clarke PM, et al. Phase I/II study of oncolytic HSV GM-CSF in combination with radiotherapy and cisplatin in untreated stage III/IV squamous cell cancer of the head and neck. Clinical cancer research : an official journal of the American Association for Cancer Research. 2010;16(15):4005-15.
19. Chang KJ, Senzer NN, Binmoeller K, Goldsweig H, Coffin R. Phase I dose-escalation study of talimogene laherparepvec (T-VEC) for advanced pancreatic cancer (ca). Journal of Clinical Oncology. 2012;30(15_suppl):e14546-e.
20. Liu BL, Robinson M, Han ZQ, Branston RH, English C, Reay P, et al. ICP34.5 deleted herpes simplex virus with enhanced oncolytic, immune stimulating, and anti-tumour properties. Gene Ther. 2003;10(4):292-303.
21. Piasecki J, Tiep L, Zhou J, Beers C. Talilmogene Iaherparepvec generates systemic T-cell-mediated anti-tumor immunity. Journal for ImmunoTherapy of Cancer. 2013;1(1):P198.
22. Cooke K, Rottman J, Zhan J, Mitchell P, Ikotun O, Yerby B, et al. Oncovex MGM-CSF – mediated regression of contralateral (non-injected) tumors in the A20 murine lymphoma model does not involve direct viral oncolysis. Journal for Immunotherapy of Cancer. 2015;3(Suppl 2):P336-P.
23. Piasecki J, le T, Ponce R, Beers C. Abstract 258: Talilmogene laherparepvec increases the anti-tumor efficacy of the anti-PD-1 immune checkpoint blockade. Cancer Research. 2015;75(15 Supplement):258-.
24. Cooke K, Estrada J, Zhan J, Mitchell P, Bulliard Y, Beltran PJ. Abstract 2351: Development of a B16F10 cell line expressing mNectin1 to study the activity of OncoVEXmGM-CSF in murine syngeneic melanoma models. Cancer Research. 2016;76(14 Supplement):2351-.
25. Louie RJ, Perez MC, Jajja MR, Sun J, Collichio F, Delman KA, et al. Real-World Outcomes of Talimogene Laherparepvec Therapy: A Multi-Institutional Experience. J Am Coll Surg. 2019;228(4):644-9.
26. National Comprehensive Cancer Center. Guidelines for Patients: Melanoma2018.
27. US Food and Drug Administration. IMLYGIC (talimogene laherparepvec) 2015 [Available from: https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/imlygic-talimogene-laherparepvec.
28. European Medicines Agency. Assessment report Imlygic International non-proprietary name: talimogene laherparepvec 2015 [Available from:

https://www.ema.europa.eu/en/documents/assessment-report/imlygic-epar-public-assessment-report_en.pdf.

29. Hirst JA, Howick J, Aronson JK, Roberts N, Perera R, Koshiaris C, et al. The need for randomization in animal trials: an overview of systematic reviews. PLoS One. 2014;9(6):e98856.
30. Macleod MR, van der Worp HB, Sena ES, Howells DW, Dirnagl U, Donnan GA. Evidence for the efficacy of NXY-059 in experimental focal cerebral ischaemia is confounded by study quality. Stroke. 2008;39(10):2824-9.
31. Landis SC, Amara SG, Asadullah K, Austin CP, Blumenstein R, Bradley EW, et al. A call for transparent reporting to optimize the predictive value of preclinical research. Nature. 2012;490(7419):187-91.
32. Zarin DA, Tse T. Sharing Individual Participant Data (IPD) within the Context of the Trial Reporting System (TRS). PLoS Med. 2016;13(1):e1001946.
33. Hodkinson A, Kirkham JJ, Tudur-Smith C, Gamble C. Reporting of harms data in RCTs: a systematic review of empirical assessments against the CONSORT harms extension. BMJ Open. 2013;3(9):e003436.
34. Rodgers MA, Brown JV, Heirs MK, Higgins JP, Mannion RJ, Simmonds MC, et al. Reporting of industry funded study outcome data: comparison of confidential and published data on the safety and effectiveness of rhBMP-2 for spinal fusion. BMJ. 2013;346:f3981.
35. Jureidini JN, McHenry LB, Mansfield PR. Clinical trials and drug promotion: selective reporting of study 329. International Journal of Risk & Safety in Medicine. 2008;20(1-2):73-81.
36. Ioannidis JP, Lau J. Improving safety reporting from randomised trials. Drug Saf. 2002;25(2):77-84.
37. Lineberry N, Berlin JA, Mansi B, Glasser S, Berkwits M, Klem C, et al. Recommendations to improve adverse event reporting in clinical trial publications: a joint pharmaceutical industry/journal editor perspective. BMJ. 2016;355:i5078.
38. Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. Nature. 2012;483(7391):531-3.
39. Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. Nature. 2014;505(7485):612-3.
40. Lalu MM, Sullivan KJ, Mei SH, Moher D, Straus A, Fergusson DA, et al. Evaluating mesenchymal stem cell therapy for sepsis with preclinical meta-analyses prior to initiating a first-in-human trial. Elife. 2016;5.
41. Llovera G, Hofmann K, Roth S, Salas-Perdomo A, Ferrer-Ferrer M, Perego C, et al. Results of a preclinical randomized controlled multicenter trial (pRCT): Anti-CD49d treatment for acute brain ischemia. Science translational medicine. 2015;7(299):299ra121.

Figure Legends

Figure 1. Study selection flow diagram

Figure 2. Preclinical and clinical efficacy of T-VEC. Four preclinical studies using mice demonstrated efficacy rates of 20-100% and clinical studies (3 melanoma and 2 mixed malignancy studies) demonstrated efficacy rates from 0-23.5%. Non-melanoma/mixed studies are represented by blue bars whereas melanoma studies are represented by orange bars. Where possible, complete regression rates of contralateral tumors for mice were used in preclinical studies and complete response was used for clinical studies. Efficacy rates decrease in the preclinical to clinical translation and upon more rigorous study design in later phase clinical trials. If possible, error bars are plotted and represent 95% confidence intervals. Some studies were not included in this analysis due to no reporting of outcome for CR.

Table 1A. *Study characteristics of included preclinical studies of T-VEC.*

Preclinical Study	Treatment	Total Number of Animals Used	Type of Cancer/Model	Efficacy Measures*	Risk of Bias (/9**)
Liu, 2003[20]	T-VEC; HSV1 wildtype immunization	90	Lymphoma (A20 murine lymphoma mouse model)	CR: 100% (n=10) (injected)	9
Piasecki, 2013[21]	T-VEC	NR	Lymphoma (A20 murine lymphoma mouse model)	CR: 70-100% of injected, 50-60% of contralateral	9
Piasecki, 2015[23]	T-VEC + Anti-PD-1	NR	Colorectal (MC-38 colon carcinoma mouse model)	CR: 80.0% (44.2-96.5%) (injected) n=10 CR: 20.0% (3.5-55.8%) (contralateral) n=10	9
Cooke, 2015[22]	T-VEC	40	Lymphoma (A20 murine lymphoma mouse model)	CR: 100% (65.5-100%) (injected) n=10 CR: 50% (23.7-76.3%) (contralateral)	9
Cooke, 2016[24]	T-VEC	20	Melanoma (B16F10 melanoma model)	NR – statistically significant tumor reduction and survival noted.	9

Table 2B. *Study characteristics of included clinical studies of T-VEC.*

Clinical Study	Treatment	Total N	Type of Cancer	Efficacy Measures*	Risk of Bias (/9**)
Hu, 2006[14] *Non-controlled Phase I*	T-VEC	30 (9 melanoma)	Breast, colorectal, melanoma, head and neck	CR: 0% (0-14.1%) PR: 0% (0-14.1%)	7
Senzer, 2009[15] *Non-controlled Phase II*	T-VEC	50	Melanoma	OR: 26.0% (15.1-40.6%) CR: 20.0% (10.5-34.1%) PR: 10.0% (3.7-22.6%)	7
Harrington, 2010[18] *Phase I/II*	T-VEC + cisplatin	17	Head and neck	CR: 23.5% (7.8-50.2%) PR: 58.8% (33.5-80.6%) OR: 82.4% (55.8-95.3%)	6
Chang, 2012[19] *Phase I*	T-VEC	17	Pancreatic	OR: 0% (0-22.9%)	6
Andtbacka, 2015[13] *Phase III*	T-VEC	295	Melanoma	DR: 16.3% (12.1-20.5%) OR: 26.4% (21.4-31.5%) PR: 15.6% (11.7-20.3%) CR: 10.8% (7.6-15.1%)	3
	GM-CSF (control)	141		DR: 2.1% (0-4.5%) OR: 5.7% (1.9-9.5%) PR: 5.0% (1.3-8.5%) CR: < 1%	
Long, 2015[16] *Phase Ib*	T-VEC + pembrolizumab	21	Melanoma	-	6
Puzanov, 2016[17] *Phase Ib*	T-VEC + IPI	18	Melanoma	DR: 44.4% (22.4-68.7%) OR: 50.0% (29.0-70.9%) CR: 22.2% (7.4-48.1%) PR: 27.8% (10.7-53.6%)	6

* DR – durable response; OR – objective response; CR – complete response/complete regression; PR – partial response; DR/OR/CR/PR definitions were based on RECIST guidelines for clinical studies. **Total number of domains that were assessed a score of high risk or unclear (maximum = 9). The nine domains include randomization, allocation concealment, blinding of participants/personnel, outcome assessment blinding, incomplete outcome reporting, selective outcome reporting, reported conflicts of interest, sample size calculation, funding.

Table 3. Construct validity assessment for preclinical studies

Author, Year	Adult Used	Animals with Advanced Stage Disease	Animals Immune to HSV	Xenograft Model Used	Used Model with a Humanized Immune System
Cooke, 2016	Unclear	No	Unclear	No	Unclear
Cooke, 2015	Unclear	No	Unclear	No	Unclear
Piasecki, 2015	Unclear	Unclear	Unclear	No	Unclear
Piasecki, 2013	Unclear	Unclear	Unclear	No	Unclear
Liu, 2003	Unclear	No	Yes	No	Unclear

Table 4. Risk of bias assessment for preclinical studies

Author, Year	Random Sequence Generation	Allocation Concealment	Blinding of Personnel	Blinded Outcome Assessment	Incomplete Outcomes Addressed	Selective Outcome Reporting	Conflicts of Interest	A Priori Sample Size Calculation	Funding
Cooke, 2016	High Risk	High Risk	Unclear	Unclear	High Risk	High Risk	Unclear	Unclear	High Risk
Cooke, 2015	High Risk	High Risk	Unclear	Unclear	High Risk	High Risk	Unclear	Unclear	High Risk
Piasecki, 2015	Unclear	Unclear	Unclear	Unclear	High Risk	High Risk	Unclear	Unclear	High Risk
Piasecki, 2013	Unclear	Unclear	Unclear	Unclear	High Risk	High Risk	Unclear	Unclear	High Risk
Liu, 2003	Unclear	Unclear	Unclear	Unclear	High Risk	Unclear	Unclear	Unclear	High Risk

Table 5. Risk of bias assessment for clinical studies

Author, Year	Random Sequence Generation	Allocation Concealment	Blinding of Participants and Personnel	Blinding of Outcome Assessors	Incomplete Outcome Data Addressed	Selective Reporting	Conflicts of Interest	Funding	Sample Size Calculation
Andtbacka, 2015	Low Risk	Low Risk	High Risk	Low Risk	Low Risk	Low Risk	High Risk	High Risk	Low Risk
Long, 2015	High Risk	N/A	High Risk	Unclear	Low Risk	Low Risk	High Risk	High Risk	Unclear
Puzanov, 2016	High Risk	N/A	High Risk	Unclear	Low Risk	Low Risk	High Risk	High Risk	Unclear
Chang, 2012	High Risk	N/A	High Risk	Unclear	Low Risk	Low Risk	High Risk	High Risk	Unclear
Harrington, 2010	High Risk	N/A	Unclear	High Risk	Low Risk	Low Risk	High Risk	High Risk	Unclear
Senzer, 2009	High Risk	N/A	High Risk	Unclear	Low Risk	High Risk	High Risk	High Risk	Unclear
Hu, 2006	High Risk	N/A	Unclear	Unclear	Low Risk	Unclear	High Risk	High Risk	Unclear

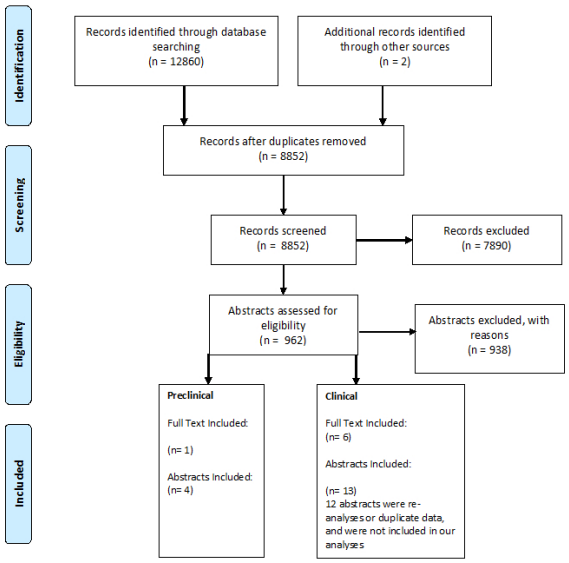

Figure 1. Study selection flow diagram

304x209mm (96 x 96 DPI)

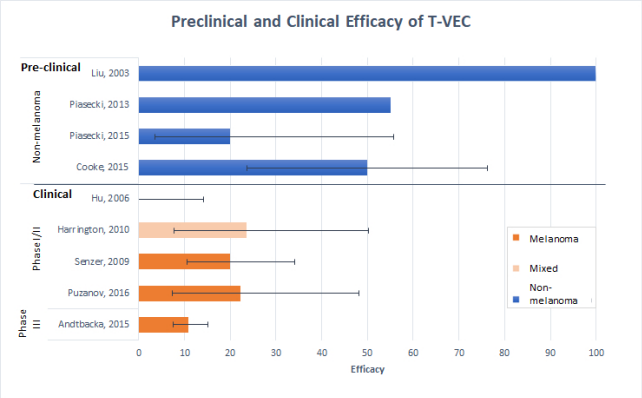

Figure 2. Preclinical and clinical efficacy of T-VEC. Four preclinical studies using mice demonstrated efficacy rates of 20-100% and clinical studies (3 melanoma and 2 mixed malignancy studies) demonstrated efficacy rates from 0-23.5%. Non-melanoma/mixed studies are represented by blue bars whereas melanoma studies are represented by orange bars. Where possible, complete regression rates of contralateral tumors for mice were used in preclinical studies and complete response was used for clinical studies. Efficacy rates decrease in the preclinical to clinical translation and upon more rigorous study design in later phase clinical trials. If possible, error bars are plotted and represent 95% confidence intervals. Some studies were not included in this analysis due to no reporting of outcome for CR.

304x209mm (96 x 96 DPI)

Online Supplementary File 1: Search Strategy

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>
Search Strategy:
---

1 Talimogen* laherparepvec.mp. (28)
2 t vec.mp. (24)
3 OncoVEX*.mp. (15)
4 Imlygic.mp. (2)
5 JS1 34*.tw. (5)
6 or/1-5 (51)
7 Oncolytic Virotherapy/ or Oncolytic Viruses/ or cancer vaccines/tu (7016)
8 (cancer adj2 (vaccine* or virus* or virotherap* or viral therap*)).tw. (5206)
9 exp neoplasms/ or cancer.tw. (3097511)
10 or/7-9 (3097684)
11 simplexvirus/ or herpesvirus 1, human/ or Herpes Simplex/ (32955)
12 (hsv1 or hsv or herpesvirus or Herpes).tw. (72684)
13 11 or 12 (77360)
14 10 and 13 (12929)
15 ((oncolyt* or cancer or tumor or tumour) adj3 (hsv1 or hsv or hsv or herpesvirus or Herpes)).tw. (846)
16 (oncolyt* adj3 (virotherap* or virus* or viral therap*)).tw. (2183)
17 or/14-16 (14671)
18 exp animal experimentation/ or exp models, animal/ or animals/ or mammals/ or vertebrates/ or exp fishes/ or exp amphibia/ or exp reptiles/ or exp birds/ or exp hyraxes/ or exp marsupialia/ or exp monotremata/ or exp scandentia/ or exp chiroptera/ or exp carnivora/ or exp cetacea/ or exp Xenarthra/ or exp elephants/ or exp insectivora/ or exp lagomorpha/ or exp rodentia/ or exp sirenia/ or exp Perissodactyla/ or primates/ or exp strepsirhini/ or haplorhini/ or exp tarsii/ or exp platyrrhini/ or catarrhini/ or exp cercopithecidae/ or gorilla gorilla/ or pan paniscus/ or pan troglodytes/ or exp pongo/ or exp hylobatidae/ or hominidae/ (5893175)
19 (animal$1 or chordata or vertebrate* or fish$2 or amphibian* or amphibium* or reptile$1 or bird$1 or mammal* or dog or dogs or canine$1 or cat or cats or hyrax* or marsupial* or monotrem* or scandentia or bat or bats or carnivor* or cetacea or edentata* or elephant* or insect or insects or insectivore or lagomorph* or rodent$2 or mouse or mice or murine or murinae or muridae or rat or rats or pig or pigs or piglet$1 or swine or rabbit$1 or sheep$1 or goat$1 or horse$1 or equus or cow or cows or cattle or calf or calves or bovine or sirenia or ungulate$1 or primate$1 or prosimian* or haplorhini* or tarsiiform* or simian*or platyrrhini or catarrhini or cercopithecidae or ape or apes or hylobatidae or hominid* or chimpanzee* or gorilla* or orangutan* or monkey or monkeys or ape or apes).tw. (4149175)
20 (preclinic$ or pre clinic$).tw. (72274)
21 or/18-20 (6474971)
22 17 and 21 (6511)
23 6 or 22 (6545)

Online Supplementary File 2. Preclinical Study Characteristics

Author, Year	Year Study Conducted	Country	Study Design	Species	Strain	Model	Type of Cancer	Baseline Tumor Size	Gender	Mean Age	Mean Weight	Co-Interventions	Duration of Follow Up
Cooke, 2015	---	USA	Interventional; Non-Controlled	Mouse	Balb/c	A20 Murine Lymphoma	Lymphoma	150 mm^3	Female	---	---	N/A	---
Piasecki, 2015	---	---	Controlled Comparison	Mouse	C57Bl/6	Syngeneic MC-38 Colon Carcinoma	Colon Cancer	---	---	---	---	Anti-PD-1	---
Piasecki, 2013	---	---	Controlled Comparison	Mouse	---	A20 Syngeneic Contralateral Model	Lymphoma	---	---	---	---	N/A	10 days
Liu, 2003	2002	UK	Controlled Comparison	Mouse	Balb/c	Syngeneic A20 Lymphoma	Lymphoma	0.5 cm diameter	---	---	---	Immunization wild type HSV1	35 days

---: not reported
N/A: not applicable
Cooke, 2016 did not provide any relevant information

Online Supplementary File 3. Preclinical Intervention and Comparator Characteristics

Author, Year	Experiment	Group	N	Dose 1	Frequency Dose 1	Duration Dose 1	Dose 2	Frequency Dose 2	Duration Dose 2	Dose 3	Frequency Dose 3	Duration Dose 3
Cooke, 2015	1	Cohort 1: TVEC	10	$3x10^4$ PFU	---	---	N/A	N/A	N/A	N/A	N/A	N/A
		Cohort 2:	10	---	---	---	N/A	N/A	N/A	N/A	N/A	N/A
		Cohort 3: TVEC	10	$3x10^6$ PFU	Every 3 days	1 week	N/A	N/A	N/A	N/A	N/A	N/A
		Cohort 4:	10	---	---	---	N/A	N/A	N/A	N/A	N/A	N/A
Piasecki, 2015	1	Int: OncoVEXmuGM-CSF + Anti-PD1	---	T-VEC: ---	Every 3 days	3 doses	Anti-PD-1: ---	Twice per wk	---	N/A	N/A	N/A
		Int: OncoVEXmuGM-CSF	---	T-VEC: ---	---	---	---	---	---	---	---	---
		Con: Anti-PD-1	---	Anti-PD-1: ---	---	---	---	---	---	---	---	---
Piasecki, 2013	1	Int: T-VEC	---	$5x10^6$ PFU	single	single	N/A	N/A	N/A	N/A	N/A	N/A
		Con: Vehicle	---	---	single	single	N/A	N/A	N/A	N/A	N/A	N/A
Liu, 2003	1	Int: JS1/34.5-/47-/mGM-CSF	10	10^6 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
			10	10^7 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
			10	10^8 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
		Int: JS1/34.5-/47-	10	10^6 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
			10	10^7 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
			10	10^8 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
		Con: vehicle	10	50µl	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
	2	Int: JS1/34.5-/47-/mGM-CSF	10	10^8 PFU/ml (50µl)	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A
		Con: Vehicle	10	50µl	Every other day	3 doses – 5 days	N/A	N/A	N/A	N/A	N/A	N/A

All doses of T-VEC were given by injection intratumorally
---: not reported
Int: intervention; Con: control; wk: week; PFU: plaque forming units
Cooke, 2016 did not provide any relevant information

Online Supplementary File 4. Clinical Study Characteristics

Author, Year	Country	Year Study Conducted	Study Type	Type of Cancer	Primary Outcomes	Secondary Outcomes
Andtbacka, 2015	USA, UK, Canada and South Africa	2009-2014	Interventional; Randomized (OPTiM Trial)	Melanoma	Efficacy: DRR	Efficacy: ORR, OS, Best Overall Response, Onset and duration of response; Time to treatment failure
Long 2015	USA, Australia, Switzerland, Spain	2014-2022	Interventional: Non-randomized, No control	Melanoma	Safety: Dose Limiting Toxicities	Efficacy: DRR, OS, Progression Free Survival; Safety: AEs
Puzanov, 2015	USA	2013-2014	Interventional: Non-Randomized	Melanoma	Safety: Dose Limiting Toxicities	Efficacy: ORR; Safety: Grade ≥3 AEs
Chang, 2012	USA	2006-2008	Interventional: Non-randomized, No control	Pancreatic Cancer	Efficacy: Detection of T-VEC in blood and urine, Presence of Anti-HSV1 Antibodies; Safety: AEs	Efficacy: ORR, Change in sum of longest tumor diameter; Change in pain intensity
Harrington, 2010	UK	2005-2010	Interventional: Non-randomized, No control	Squamous Cell Carcinoma of the head and neck	Safety: AEs	Efficacy: Antitumor Activity, OS*, Complete Response*, Partial Response*; Progression Free Survival*
Senzer 2009	USA	2005-2008	Interventional; (non-controlled, non randomized)	Melanoma	Efficacy: ORR	Efficacy: OS; Safety: AEs
Hu, 2006	USA, UK	---	Interventional: Non-randomized, No control	Breast, Colorectal, Melanoma, Head and Neck	Efficacy: Biodistribution; Safety: AEs	Efficacy: GM-CSF expression, HSV antigen associated necrosis, Viral Replication; Local Reactions

---: Not Reported
*: not reported a priori
AEs - adverse events; CHN– cutaneous head and neck; DRR – durable response rate; ECOG – Eastern Cooperative Oncology Group; ORR- objective response rate; OS – overall survival

Online Supplementary File 5. Clinical Patient Characteristics

Author, Year	Group	Patients (N)	Median Age (range)	Sex (n; F)	Metastasis Stage (n; Stage IVM1b/c)	Line of Therapy (n; first line)	HSV Serostatus (n; Seropositive, n; unknown)
Andtbacka, 2015	T-VEC	295	63 (22-94)	122	131	138	175, 23
Long, 2015	T-VEC + Pembrolizumab	21	58	13	11	---	---
Puzanov, 2015	T-VEC + Ipilimumab	18	---	---	---	18	---
Chang, 2012	T-VEC	17	54	6	---	---	---
Harrington, 2010	T-VEC and Chemo radiotherapy	17	58 (41-74)	2	3	---	---
Senzer, 2009	T-VEC	50	62 (34-88)	28	24	0	36, 1
Hu, 2006	T-VEC	30	55 (30-80)	23	---	0	19

--- : Not Reported

Online Supplementary File 6. Clinical Intervention and Comparator Characteristics

Author, Year	Arm	Dose 1	Time of Dose 1	Frequency of Dose 1	Dose 2	Time of Dose 2	Frequency of Dose 2	Dose 3	Time of Dose 3	Frequency of Dose 3	Intervention Window	Follow Up Duration
Andtbacka, 2015	T-VEC	10^6 PFU/ml (≤4ml)	Week 1	single	10^8 PFU/ml (≤4ml)	Week 4	Q2W	N/A	N/A	N/A	Median: 23 wks (0.1-79 wks)	Median: 44 mo (32-58 mo)
	GM-CSF	125 µg/m^2	Week 1	Once daily 14/28 day cycles	N/A	N/A	N/A	N/A	N/A	N/A	Median: 10 wks (0.6 to 72 wks)	---
Long, 2015	T-VEC + Pemb.	TVEC: 10^6 PFU/ml	Day 1	single	T-VEC: 10^8 PFU/ml	Day 22	Q2W	Pemb: 200 mg	Day 36	Q2W	Median TVEC: 13 wks Median Pemb: 10 wks	---
Puzanov, 2015	T-VEC + Ipilimumab	TVEC: 10^6 PFU/ml (≤4ml)	Week 1	single	TVEC: 10^8 PFU/ml (≤4ml)	Week 4	Q2W	Ipilimumab: 3mg/kg	Week 6	Q3W	TVEC: until DLT Ipi: 12 wks	17 mo minimum
Chang, 2012	Cohort 1	10^4 PFU/ml	Week 1*	single	10^5 PFU/ml	Week 4*	Q3W*	N/A	N/A	N/A	up to 15 wks	---
	Cohort 2	10^5 PFU/ml	Week 1	single	10^6 PFU/ml	Week 4	Q3W	N/A	N/A	N/A	up to 15 wks	---
	Cohort 3	10^6 PFU/ml	Week 1	single	10^7 PFU/ml	Week 4	Q3W	N/A	N/A	N/A	up to 15 wks	---
Harrington, 2010	Cohort 1	T-VEC: 10^6 PFU/ml	Day 1	Q3W	Cisplatin: 100mg/m^2	Day 1	Q3W	N/A	N/A	N/A	Up to 9 weeks	Median: 29mo (19-40mo)
	Cohort 2	T-VEC: 10^6 PFU/ml	Day 1	single	T-VEC: 10^7 PFU/ml	Day 22	Q3W	Cisplatin: 100mg/m^2	Day 1	Q3W	Up to 9 weeks	Median: 29mo (19-40mo)
	Cohort 3	T-VEC: 10^6 PFU/ml	Day 1	single	T-VEC: 10^8 PFU/ml	Day 22	Q3W	Cisplatin: 100mg/m^2	Day 1	Q3W	Up to 9 weeks	Median: 29mo (19-40mo)
	Cohort 4	T-VEC: 10^6 PFU/ml	Day 1	single	T-VEC: 10^8 PFU/ml	Day 22	Q3W	Cisplatin: 100mg/m^2	Day 1	Q3W	Up to 9 weeks	Median: 29mo (19-40mo)
Senzer, 2009	T-VEC	10^6 PFU/ml (≤4ml)	Week 1	Single	10^8 PFU/ml (≤4ml)	Week 4	Q2W	N/A	N/A	N/A	Max: 48 wks Median: 11 wks	Median: 18 mo (11-36 mo)
Hu, 2006	Single Dose Group 1	10^6 PFU/ml	---	single	N/A	N/A	N/A	N/A	N/A	N/A	Single dose	6 wks
	Single Dose Group 2	10^7 PFU/ml	---	single	N/A	N/A	N/A	N/A	N/A	N/A	Single dose	6 wks
	Single Dose Group 3	10^8 PFU/ml	---	single	N/A	N/A	N/A	N/A	N/A	N/A	Single dose	6 wks
	Multi-dose Group 1	10^6 PFU/ml	---	single	10^7 PFU/ml	---	Q1-3W	N/A	N/A	N/A	3-9 wks*	6 wks post final injection
	Multi-dose Group 2	10^6 PFU/ml	---	single	10^8 PFU/ml	---	Q1-3W	N/A	N/A	N/A	3-9 wks*	6 wks post final injection

Author, Year	Arm	Dose 1	Time of Dose 1	Frequency of Dose 1	Dose 2	Time of Dose 2	Frequency of Dose 2	Dose 3	Time of Dose 3	Frequency of Dose 3	Intervention Window	Follow Up Duration
	Multi-dose Group 3	10^8 PFU/ml	---	Q1-3W	N/A	N/A	N/A	N/A	N/A	N/A	3-9 wks*	6wks post final injection

DLT: Dose Limiting Toxicity; Pemb: Pembrolizumab; Q2W: every two weeks Q3W: every three weeks; Q1-3W: every 1-3 weeks; Q6W every 6 weeks
--- : not reported
T-VEC was given by intra-tumoral injection in all studies

Online Supplementary File 7. Preclinical Efficacy Data

Author, Year	Experiment	Group	N – Animals Studied	N – Lesions Studied	Baseline Mean Tumor Measure (Standard Error of Mean)	Final Mean Tumor Measure (Standard Error of Mean)	CR - Injected	CR - Contralateral	Duration of Follow Up
Cooke, 2015	1	INT: TVEC $3x10^4$ PFU	10	---	---	---	---	---	---
		INT: TVEC $3x10^6$ PFU	10	10	~150mm^3	---	10/10	5/10	---
		---	10	---	---	---	---	---	---
		---	10	---	---	---	---	---	---
Piasecki, 2015	1	INT: OncoVexmuGM-CSF	---	---	---	---	---	---	---
		INT: OncoVexmuGM-CSF + Anti-PD-1	---	20	---	---	8/10	2/10	---
		CON: Anti Pd-1	---	---	---	---	---	---	---
Piasecki, 2013	1	INT: T-VEC	---	---	---	---	70-100%	50-60%	10 days
		CON: Vehicle	---	---	---	---	---	---	---
Liu, 2003	1	INT: JS1/34.5-/47-/mGM-CSF; injected	10	N/A	5.2mm (0.34)	0.004mm (0.31)	N/A	N/A	22 days
		INT: JS1/34.5-/47-/mGM-CSF; uninjected	10 (same as injected)	N/A	5.7mm (0.29)	1.1 mm (0.73)	N/A	N/A	22 days
		INT: JS1/34.5-/47-; injected	10	N/A	5.4mm (0.37)	1.4 mm (1.36)	N/A	N/A	22 days
		INT: JS1/34.5-/47-; uninjected	10 (same as injected)	N/A	6.2mm (0.29)	5.4mm (2.01)	N/A	N/A	22 days
		CON: Vehicle; injected	10	N/A	5.4mm (0.40)	11.9mm (2.69)	N/A	N/A	22 days
		CON: Vehicle; uninjected	10 (same as injected)	N/A	5.6mm (0.46)	13.2mm (2.76)	N/A	N/A	22 days
	2	INT: JS1/34.5-/47-/mGM-CSF	10	N/A	5.5mm (0.34)	2.2mm (1.6)	N/A	N/A	21 days
		CON: Vehicle	10	N/A	5.6mm (0.23)	13.8mm (1.2)	N/A	N/A	21 days

--- : not reported
INT: intervention
CON: control
Liu 2003 data from experiment 1 taken from 10^8 dose
Cooke, 2016 did not provide any relevant information

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	4
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5-6
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	5-6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	7

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	6
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	7-8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9-10
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	8-9
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	8-9
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	9-10
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11-12
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	12-13
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	13
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	14

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.