Validation of a classification system for treatment-related mortality in children with cancer

Hadeel Hassan,[1,2] Menie Rompola,[1,2] Adam Woolf Glaser,[1,2]
Sally Elizabeth Kinsey,[1,2] Robert Stephen Phillips[1,3]

[1]Department of Paediatric Haematology and Oncology, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[2]Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, UK
[3]Centre for Reviews and Dissemination, University of York, York, UK

Correspondence to
Dr Hadeel Hassan;
hadeelhassan@doctors.org.uk

ABSTRACT

Background Death not directly due to cancer has been termed 'treatment-related mortality' (TRM). Appreciating the differences between TRM and disease-related death is critical in directing strategies to improve supportive care, interventions delivered or disease progression. Recently, a global collaboration developed and validated a consensus-based classification tool and attribution system.

Objectives To evaluate the reliability of the newly developed consensus-based definition of TRM and explore the use of the cause-of-death attribution system outside the centre it was initially validated (Toronto, Canada). In the initial study, reviewers listed multiple causes of death. In this study, reviewers identified a primary cause for simplicity.

Setting The paediatric haematology and oncology department at Leeds Teaching Hospital in Leeds, UK.

Participants Two consultants and two clinical research associates (CRAs).

Methods Thirty medical records of the most recent deaths in children with cancer, 2 and 4 weeks prior to death, were anonymised and presented to the participants. Reviewers independently classified deaths as 'treatment related mortality' or 'not treatment related' according to the algorithm developed. When TRM occurred, reviewers applied the cause-of-death attribution system to identify the primary cause of death. Inter-relater reliability was assessed using the kappa statistic (k).

Main outcome Inter-relater reliability between CRA and consultants.

Results Reliability of the classification was deemed 'very good' between CRA and consultants (k=0.86, 95% CI 0.72 to 0.97). Ten deaths were classified as TRM, of which infection was the most frequent cause identified. Reviewers disagreed on the primary cause of death (eg, respiratory vs infection) when applying the cause-of-death attribution system in six cases and probable and possible causes in four cases. The study identified how the algorithm may not detect TRM in patients receiving non-curative therapy.

Conclusions The classification and cause of death attribution system could be implemented in different healthcare settings. Adaptation of the classification tool in patients receiving non-curative interventions and the cause of death attribution system should be considered.

What is already known on this topic?

► Treatment-related mortality (TRM) is poorly defined in paediatric oncology and haematology.
► A global collaboration developed and validated a consensus-based classification tool and attribution system (in Toronto, Canada).

What this study hopes to add?

► The classification tool is not designed to detect TRM in patients receiving non-curative therapy.
► The classification and cause of death attribution system can be used in different healthcare settings.

suggesting more than 80% of those diagnosed in high-income countries will survive.[2] This still means 20% of children will die—around 80 000 deaths worldwide in 2012.[2] The cause of death may be due directly to the disease, or complications or toxicities of interventions delivered. The increasing success of cancer-directed interventions has contributed to a higher relative proportion of deaths due to associated toxicities.[3] Death not directly due to the cancer has been broadly termed 'treatment-related mortality' (TRM). This includes death from infection, bleeding and organ dysfunction.[4] Incidence varies according to the underlying disease, stage of treatment and socioeconomic status.[5] Identifying the nature of the mortality is critical in helping researchers and clinicians improve survival in particular diagnostic groups by focusing on strategies to improve supportive care when death is predominantly through TRM.

Despite the importance of TRM, varying definitions have been used. Systematic reviews[6 7] identified significant heterogeneity in TRM definitions used in randomised trials. This inspired a global collaboration to develop a consensus-based classification tool for ascribing death as TRM, and further

BACKGROUND

Survival of children diagnosed with malignancies has improved with recent estimates

specific attribution of cause of death,[1] primarily to be used by clinical research associates (CRAs). Reliability for the TRM classification was 'very good' between medical and CRA reviewers that had been involved in developing the system (kappa=0.92, 95% CI 0.78 to 1.00).[1]

An ideal classification system for TRM should be applicable across different countries, treatment protocols and healthcare settings. To assess this, it is important to attempt to further validate the classification and attribution system in a different setting, with professionals uninvolved in the development of the system. This paper reports such validation at a large regional paediatric oncology centre in a different continent to the one in which the system was originally developed. In the original study, reviewers attributed death to multiple probable or possible causes. While developing the study protocol, we decided to limit the number of causes of death to one primary probable or possible cause for simplicity.

OBJECTIVES

This study aimed to evaluate the reliability of the newly developed consensus-based definition of TRM and explore the use of the cause-of-death attribution system at a regional paediatric oncology centre in Leeds, England. In the initial study, reviewers listed multiple causes of death. However, in this study, reviewers identified a primary cause for simplicity.

METHODS

Eligible patient records were those of patients treated for a malignancy or who underwent a haematopoietic stem cell transplant (HSCT) for a non-malignant diagnosis at Leeds Children's Hospital (Leeds, UK) while aged 18 years or younger at diagnosis. Five cases were excluded as patients had died following relapse after the age of 18 years or the medical records were not located. All included patients died between 2014 and 2016. Thirty patient records were included. Copies of the clinical records, with information from both 2 weeks prior to death and with the information extending back to 4 weeks prior to death, were anonymised. This resulted in 60 sets of anonymised case notes (30 patients, each with two time periods), which were presented in a different random order for each assessor. Four participants were
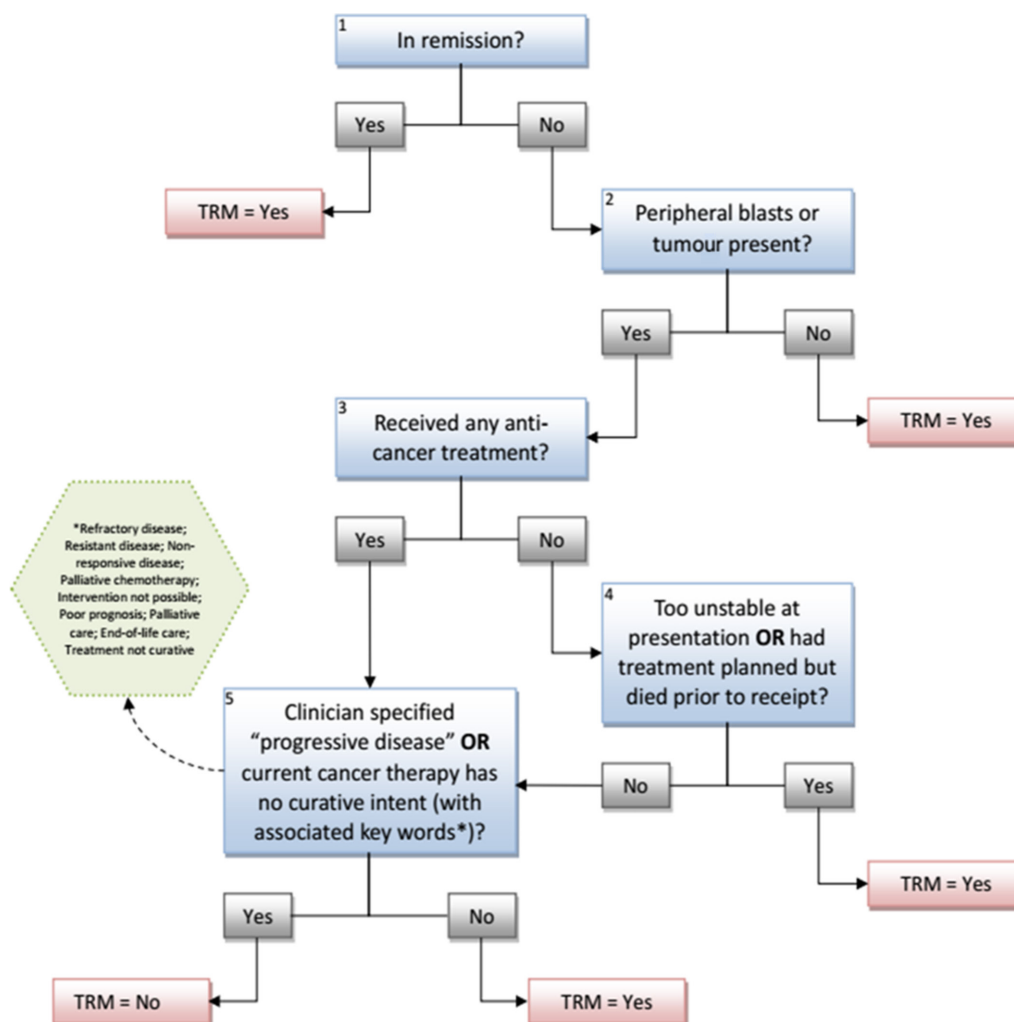
Figure 1 Classification of TRM in children with cancer, taken from.[7] TRM, treatment-related mortality.

identified to review the case notes; the two CRAs were a data analyst (AF) and research nurse (JT), and the two senior clinicians were a consultant paediatric oncologist and consultant paediatric haematologist (AWG and SEK).

The study was undertaken on a single afternoon. After agreeing and signing a consent form, participants received a 10 min educational presentation explaining how to use the system and how the study would be undertaken. The reviewers then independently classified each death according to the algorithm (figure 1). For cases assessed as TRM, the reviewers were asked to apply the cause-of-death attribution system (online supplementary file 1) to identify a primary cause of death. Following the individual completion of the assessments, a moderated group discussion was undertaken with notes recorded by two facilitators and used to supplement the themes of the discussion.

Inter-relater reliability was assessed using the kappa statistic (k). Criterion validity was assessed by assuming classification by the consultants as the gold standard. Group consensus classification between and within the CRAs and consultant group was evaluated using the Cohen's kappa statistic and across all individuals using the Fleiss' kappa statistic. The strength of agreement was defined as slight (0.00–020), fair (0.21–0.40), moderate (0.41–0.60), good (0.61–0.80) and very good (0.81–1.00).[8] A numerical code was used to combine agreement/disagreement between the individual consultants (TRM was recorded as '0' and non-TRM outcomes were recorded as '1' in Excel). When calculating inter-relater reliability between the CRAs and consultants if individual disagreement was recorded, then the outcome was recorded as '2'. Based on the previously published study,[1] we decided to include 30 cases for analysis. A sample size of 27 deaths determined whether k was good (ie, ≥0.61), with a power of 0.80, and two-sided α of 0.05 and assuming that TRM accounted for 20% of deaths.[7] A further 30 cases would be reviewed if validity was inadequate (defined a priori as k<0.6). Calculation of the k statistic was completed using the R studio irr package, and a bootstrap with 200 iterations was used to calculate 95% CIs.[9] Comparison of cause of death was qualitative, and to provide further insight, reviewers participated in a discussion of the use of the algorithm and cause of death attribution system.

RESULTS
Demographics
Age of identified patients ranged from less than 1–17 years and 57% (17) were male. Sixty-seven per cent (20) were diagnosed with solid tumours, and 33% (10) were diagnosed with malignant haematological conditions. Collectively, sarcomas were the most frequent solid tumour diagnosed (12 cases, 40% of total), followed by central nervous system (CNS) tumours (nine cases, 30% of total). Twenty-seven per cent (eight cases) of these patients had presented with metastatic disease at diagnosis. All patients with malignant haematological conditions were diagnosed with leukaemia. Forty per cent (12 cases) of patients had either received a transplant or presented with relapsed disease. A summary is enclosed in online supplementary file 2.

Classification of TRM
Ten deaths (33%) were identified as TRM by at least one reviewer. Fifteen per cent (3/20) of patients diagnosed with solid tumours, and 80% (8/10) of patients diagnosed with malignant haematological conditions who died were classified as TRM. Reliability of classification was very good between CRAs and consultants, with a kappa statistic of 0.86 (95% CI 0.72 to 0.97, with disagreement on three deaths). There was also very good agreement between CRAs (k=0.96, 95% CI 0.87 to 1.00, disagreement on one record) and between consultants (k=0.85, 95% CI 0.67 to 0.97, with disagreement on two deaths) (see table 1).

When the 2-week and 4-week data were examined, there was no difference between the assessments of each of the four assessors (see table 1).

Cause of death attribution system
Table 2 summarises the diagnoses of the 10 patients whose deaths were classified as TRM, alongside the cause of death as attribution by each assessor.

In brief, there was unanimous agreement on cause of death of three cases (J, V and AA); the other seven cases has inconsistencies in either the nature of cause of

Table 1 Summary of kappa statistic (k) and 95% CIs of independent reviewers, consultants, CRAs and between CRA and consultants for all total case reviews, 4 weeks and 2 weeks prior to death using the cause of death attribution system

Inter-rater comparison	Total k (95% CI)	4 weeks k (95% CI)	2 weeks k (95% CI)
Independent reviewers*	0.92 (0.83 to 0.98)	0.91 (0.76 to 1.00)	0.92 (0.79 to 1.00)
Consultants†	0.85 (0.67 to 0.97)	0.85 (0.59 to 1.00)	0.84 (0.59 to 1.00)
CRA†	0.96 (0.87 to 1.00)	0.85 (0.59 to 1.00)	0.85 (0.60 to 1.00)
CRA versus consultants†	0.86 (0.72 to 0.97)	0.87 (0.66 to 1.00)	0.86 (0.67 to 1.00)

*Calculated using the Fleiss' kappa statistic (between four reviewers).
†Calculated using the Cohen's kappa statistic (between two reviewers or two groups).
CRA, clinical research associates.

Table 2 Summary of cause of death attribution by reviewers for deaths classified as TRM

Case	Diagnosis	CRA 1	CRA 2	Consultant 1	Consultant 2
J	AML prior to HSCT	**Immunomediated**	**Immunomediated**	**Immunomediated**	**Immunomediated**
M	ATRT	**Acute symptomatic intracranial haemorrhage**	**Acute symptomatic intracranial haemorrhage**	Not reported	**Acute symptomatic intracranial haemorrhage**
V	AML	**Infection**	**Infection**	**Infection**	**Infection**
X	B cell ALL	**Acute symptomatic intracranial haemorrhage**	**Acute symptomatic intracranial haemorrhage**	**Infection**	Infection
Y	T cell ALL	**Respiratory**	Infection	**Acute symptomatic intracranial haemorrhage**	Infection
Z	B cell ALL	**Respiratory**	Infection	**Respiratory**	Infection
AA	B cell ALL	Infection	Infection	Infection	Infection
AB	Ependymoma	Not reported	Infection	Not reported	Infection
AC	T cell ALL	**Worsening symptomatic graft versus host disease**	**Worsening symptomatic graft versus host disease**	**Worsening symptomatic graft versus host disease**	**Acute symptomatic pulmonary haemorrhage**
AD	B cell ALL post-HSCT	Unclear	Difficult case	**Worsening symptomatic graft versus host disease**	**Acute kidney Injury**

Bold font=probable causes of death.

ALL, acute lymphoblastic leukaemia; AML, acute myeloid lukemia; ATRT, atypical teratoid rhabdoid tumour; CRA, clinical research associates; HSCT, haematopoietic stem cell transplant; TRM, treatment-related mortality.

death (eg, infection vs immune mediated) or strength of evidence (ie, probable vs possible). In one complex scenario (AD), the CRAs did not feel able to record a cause of death.

Postreview discussion

The reviewers all agreed that the algorithm was straightforward to use and that it would be beneficial to have a standardised tool and attribution system to use in trials nationally and globally.

Three difficulties with the system were identified: first, how to address the patients who may die from TRM while receiving non-curative therapy and particularly how the algorithm could be used as part of palliative care trial. In one particular case, a patient taking palliative etoposide following a diagnosis of relapsed ALL developed a febrile illness and died. Second, how should deaths in children after the completion of treatment be classified? A child (case V) who had completed treatment for standard risk AML died of overwhelming pneumococcal septicaemia 6 months after end of treatment while in complete remission. Another further case the assessors had difficulty with (case M) followed presentation acutely with signs of raised intracranial pressure and large mass noted on CT. This patient died on the operating table while receiving surgical intervention.

Attribution of cause of death was felt to be more difficult than ascribing a death as TRM or not. The case the reviewers felt was most difficult to assign was patient who died of multiorgan failure following a HSCT (AD); both CRAs independently decided to not attribute a cause of death.

DISCUSSION

This study is, to the best of our knowledge, the first revalidation of the standardised definition of TRM and cause of death attribution system for patients with paediatric cancer.[1] It demonstrates that the system is reliable and established its validity in an alternative centre and healthcare system with different treatment protocols. It can be used after very limited training, with 'very good' agreement between assessors irrespective of discipline (Fleiss' kappa 0.92, 95% CI 0.83 to 0.98). The study confirms the observations of the development group and shows that information from 2 weeks prior to the death of a patient is sufficient to consistently attribute death to TRM or disease.

Although consultants' opinions are considered gold standard, in this study we identified how even experienced clinicians may disagree on use of the algorithm. Consultants disagreed on the classification of death in two cases; this may have occurred due to the individual consultant's clinical experience or previous contact with the patients. Even though the cases were anonymised and randomised, the physicians may have recognised the patient due to their potential clinical involvement in direct patient care. The differences identified highlight how the TRM classification tool is unlikely to ever have perfect agreement between reviewers irrespective of experience, and clinical, scientific knowledge.

In this study, reviewers attributed death to one primary probable, or possible, cause. While developing the study protocol, we decided to limit the number of causes of
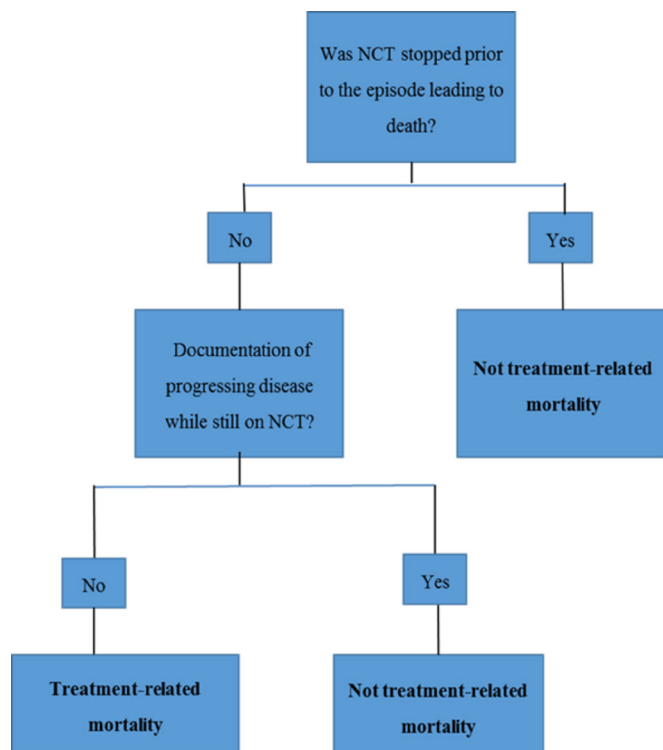
Figure 2 Proposed classification of TRM in children receiving non-curative therapy (NCT) only. TRM, treatment-related mortality.

death for simplicity. However, reviewers found it challenging to identify only one cause of death and distinguish between probable and possible causes. In the original study, reviewers attributed death to multiple probable or possible causes. Despite this difference, reviewers only agreed on the causes of death on three of nine cases classified as TRM. This is similar to the results in this study of which reviewers agreed on 3 of 10 cases.

Since the development of this study, a standard operating procedure TRM web-based tool has been published (https://www.sungresearch.com/trm-training-manual/) and includes working examples. Use of this tool when delivering the training package should help clarify how to use the cause-of-death attribution system and minimise misunderstanding. Currently, the web-based tool is available in English; having the tool available in other languages could potentially reduce confusion and improve harmonisation across clinical trials.

Our study highlights specific challenges with the system as it currently exists, both with the classification of TRM and the attribution of a specific cause of death. Fundamentally, this approach defines deaths as either 'treatment related' or 'cancer related'. This gives rise to a semantic challenge; 'treatment-related mortality' implies that deaths that come under this term occur directly because of the therapies delivered. However, the classification system classifies deaths that occur prior to commencement of anticancer therapy that are not directly attributable to the cancer (eg, tumour lysis syndrome in high-count leukaemia) as cases of TRM. This

clash of language and 'common sense' may confuse users of the classification tool, for example, case M in which a patient presented acutely with signs of raised intracranial pressure and died on the operating table.

A deeper challenge to this system addressed the philosophical distinction of assigning all deaths into one of two categories: cancer or treatment related. There is a convincing argument that a third category of death should be attributable, 'other non-cancer death', for those who die of an event or illness external to their malignancy. This problem is particularly evident if the current system is to be used after the completion of treatment. For example, a patient dies as a passenger in an air traffic accident, 4 years after treatment for a localised Wilms tumour. In the current system, this death is classified as TRM, even though the death is unrelated to the child's cancer diagnosis.

Conversely, it is also important to note how a diagnosis of cancer may be included in 'non-cancer deaths'. For example, a patient may commit suicide some years after the completion of treatment because of the psychological effects of their diagnosis or treatment. Any addition to the system would need to be sensitive to these potential issues.

We have identified the need to further refine the approach to categorising cause of death in patients with cancer receiving care without intent to cure. This is particularly important if the system is applied in 'routine' settings, assessing deaths in the palliative setting rather than in the original setting of use within a curative trial. Increasingly, individuals destined not to be cured are living for lengthier periods due to participation in clinical trials/studies. This group of patients currently have all deaths classified as 'not treatment-related mortality' as clinicians would have specified progressive disease or that cancer therapy has no curative intent. This algorithm may fail to identify a significant group of patients who may die of causes amenable to better supportive care while receiving palliative care. For example, a patient can die of overwhelming sepsis while receiving palliative etoposide for refractory neuroblastoma. This could be addressed by modifying the algorithm, for this type of use. Another proposal includes using a separate classification tool for patients on non-curative therapy trials (figure 2), although this should be further developed in conjunction with palliative care physicians and researchers. The counterargument to this suggested change is the risk of adding complexity to a simple, effective tool that can be used by people of different skills and healthcare settings globally. It will also have similar issues with interpretation as with initial algorithm.

Reviewers failed to agree on a primary cause of death in six episodes and probable and possible causes in four cases. Differences in cause of death allocated could be attributed to the reviewers' previous experience, clinical expertise and interpretation of the clinical records, particularly in light of potential previous direct clinical

involvement with the cases under review. Currently, the tool is intended for use by any CRA; however, it may require users to have a certain level of experience or clinical expertise, and agreement may be reduced with CRA who are new to the role.

Understanding and interpretation of the system as proposed for attribution of a specific mechanistic cause of death could potentially be improved by dedicating more time during the presentations and using the newly developed web-based training tool. Alternatively, the cause of death attribution system could be refined, and multiple causalities permitted.

CONCLUSIONS

We have been able to confirm the reproducibility and criterion validity of the TRM classification system. We believe this supports the hypothesis that the classification system can be implemented easily and effectively in different healthcare settings, thereby improving consistency and accuracy of outcome reporting in clinical trials. The TRM classification system will be of immense value in the evaluation of deaths in the palliative setting. We propose the addition of a separate classification tool in patients receiving palliative treatment.

REFERENCES

1. Alexander S, Pole JD, Gibson P, *et al*. Classification of treatment-related mortality in children with cancer: a systematic assessment. *Lancet Oncol* 2015;16:e604–e610.
2. CRUK. Children's cancers mortality in Europe and worldwide, 2014. http://www.cancerresearchuk.org/health-professional/cancer-statistics/childrens-cancers/mortality#heading-Three (cited 2016 September).
3. Kremer LC, Mulder RL, Oeffinger KC, *et al*. A worldwide collaboration to harmonize guidelines for the long-term follow-up of childhood and young adult cancer survivors: a report from the International Late Effects of Childhood Cancer Guideline Harmonization Group. *Pediatr Blood Cancer* 2013;60:543–9.
4. Blanco E, Beyene J, Maloney AM, *et al*. Non-relapse mortality in pediatric acute lymphoblastic leukemia: a systematic review and meta-analysis. *Leuk Lymphoma* 2012;53:878–85.
5. White Y, Castle VP, Haig A. Pediatric oncology in developing countries: challenges and solutions. *J Pediatr* 2013;162:1090–1.
6. Ethier MC, Blanco E, Lehrnbecher T, *et al*. Lack of clarity in the definition of treatment-related mortality: pediatric acute leukemia and adult acute promyelocytic leukemia as examples. *Blood* 2011;118:5080–3.
7. Tran TH, Lee M, Alexander S, *et al*. Lack of treatment-related mortality definitions in clinical trials of children, adolescents and young adults with lymphomas, solid tumors and brain tumors: a systematic review. *BMC Cancer* 2014;14:612.
8. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med* 2012;22:276–82.
9. Team R. RStudio: Integrated Development for R. *RStudio, Inc* 2015.