Managing the surge in demand for blood following mass casualty events. Early automatic restocking may preserve red cell supply

Short Title: Managing blood demand in mass casualty events

Simon Glasgow¹ BSc MD, Christos Vasilakis² MSc PhD, Zane Perkins¹ MD Nigel Tai¹ MS MD, Karim Brohi¹ MD

Affiliations

¹Centre for Trauma Sciences, Blizard Institute of Cell & Molecular Science, Queen Mary University of London, UK.

²Centre for Healthcare Innovation and Improvement, School of Management, University of Bath.

Author email addresses

simon.glasgow@bartshealth.nhs.uk
c.vasilakis@bath.ac.uk
zane.perkins@bartshealth.nhs.uk
nigel.tai@bartshealth.nhs.uk
karim.brohi@bartshealth.nhs.uk
Address for correspondence

Professor Karim Brohi
Centre for Trauma Sciences
Blizard Institute
Queen Mary University of London
Newark St
London E1 2AT
k.brohi@qmul.ac.uk
Tel: +44 (0)20 7882 2288
Fax: +44 (0)20 7882 2180

Funding Sources

Simon Glasgow was supported by a Research Fellowship award, Royal College of Surgeons of England, 35-43 Lincoln’s Inn Fields, London, WC2A 3PE.

Conflicts of Interest

The authors declare no conflicts of interest in relation to this study.

Author Contribution

All authors were involved in the design, analysis and writing of the paper.
Background

Blood is a critical resource for responding to mass casualty events (MCEs) (1). Individual trauma patients have been known to consume as much as a 100U of red blood cells (RBCs) (2-4). Such cases are manageable at well-equipped trauma centres (TCs), however, rates of civilian MCE massive transfusion (MT) of over 10 units of RBCs to single casualties can be as high as 10% (5, 6). These MT rates are similar to warfare and civilian injuries from modern events are increasingly resembling that of the battlefield, with large-scale shootings using automatic weapons reported in India, Norway and France over recent years (7-11). The surge in RBC demand from casualties requiring a MT has the potential to decimate an individual hospital’s blood bank inventory, severely compromising a unit’s ability to treat bleeding casualties effectively (12-14). Strategies for best managing this evolving demand are crucial to reduce the in-hospital preventable mortality following these events.

There has been a nine-fold increase in global mortality from terrorist incidents since the year 2000 (15, 16). This has led to the term MCE becoming increasingly synonymous with terrorism and a renewed interest in MCE transfusion planning (6, 14, 17-20). Modelling methodologies have been applied previously in the context of MCEs, predominantly within the remit of pre-hospital care (21-28). Examples of in-hospital event modelling have indicated the perceived capacity of hospitals to manage MCEs to be overly-optimistic, one specifically identified blood product volume as a limiting factor in the event response (29, 30). In addition, the provision of RBCs in MCEs is time critical. The majority of all trauma deaths in the first hour and approximately 50% of deaths in the first 24 hours are due to bleeding (31, 32). The result is two
thirds of an event’s total RBC requirement on the first day is consumed within the first four hours (14, 33, 34). Both time and supply elements must be overcome together if MCE RBC provision is to be managed effectively.

The overall objective of this study was to investigate strategies for optimizing in-hospital RBC provision to casualties in MCEs. This was approached through the application of computer simulation modelling to investigate two related aims. The first aim was to establish the effect of varying on-shelf RBC stock hold at a TC on the ability to adequately treat increasing numbers of casualties with RBCs. The second aim was to determine the effect time of RBC supply restocking has on casualty treatment rates compared to maintaining the equivalent stock levels on-shelf at a TC.

Methods

A computerised discrete event simulation (DES) model was constructed in Arena Simulation Enterprise Suite version 14.0 (Rockwell Automation, Pittsburgh, USA). DES modelling is recommended for models involving determination of scarce resource allocation and queue minimisation problems (35). Computer simulation involves the execution of repeated runs of a mathematical model to provide a series of model solutions, which collated as a set of statistics can be interpreted to measure system performance (36). The discrete nature refers to the progression of the simulation clock only when an event which changes the state of one of the model variables occurs.
The model setting was a generic UK TC under the surge conditions of an unspecified MCE during the first 72 hours following an event (when transfusion pressure is greatest). Within the model only those procedures involved in the provision of RBCs to casualties from arrival through to receipt of their complete designated RBC demand were considered. No other blood products were considered in this RBC specific model. The TC itself represented the boundaries of the model which was divided into three distinct areas of operation: casualty assessment, casualty treatment and the transfusion laboratory. A schematic of the model design is presented in Figure 1.

On initiation of the simulation the first casualty arrives at the TC which begins in a zero state. As casualties arrive they are triaged as either Priority one (P1) or two (P2). The priority system is specific to MCEs and allows a clear distinction between critically injured P1 casualties at risk of major haemorrhage and those who are not, but who may still require RBCs. Priority three (P3) walking wounded casualties were not included in the model as they would not be expected to require blood. A list of model assumptions applied to manage system complexity is provided in Table 1.

Following triage, casualties flow through the model undergoing assessment for blood loss, assignment of an overall RBC demand and finally the provision of this demand. Provision of RBC is initially performed through the transfusion of emergency type O universal donor RBCs, followed by type specific RBCs when their individual blood group has been determined.
Casualty blood grouping is performed via the transfusion laboratory section of the model in parallel to their receiving emergency transfusions. Casualties exit the model if they not to require a transfusion of if their transfusion demand has been fully met.

The model structure and input data required to drive it were derived from a number of sources, including: a literature review of 100 years of historical MCEs (previously published), direct observation of real-world TC transfusion laboratory processes and interrogation of both civilian and military trauma databases (37). Throughout a simulation Arena refers directly to the input values or samples from a probability distribution provided for each model variable when making a decision or executing an event. These data inputs were all defined during the model’s construction and are outlined in Table 2.

Model data was collected throughout the simulation run to measure the system performance through two principal outcome measures. The first was the percentage of bleeding P1 and P2 casualties receiving their assigned RBC demand within the time allocated by their triage category (within one hour for P1s and four hours for P2s). The second outcome measure was the median time point at which emergency type O RBC inventory levels were exhausted (a surrogate for the inability to treat and therefore receive further bleeding casualties).

Prior to experimentation, a full model evaluation study was performed (38). The model was verified through expert review of the source code and visual inspection of the design. Model validation was performed through comparison with data from the 2005 London bombings - the
largest UK MCE in recent years. This included 27 P1 and P2 casualties and the transfusion of 164 units (U) of RBC. Sensitivity analysis of all input variables identified the need for 100 replications of the simulation for each experimental scenario for suitably accurate results. As we wished to experiment across a range of MCE sizes, for each scenario, the simulation run was repeated 100 times across increasing casualty loads from 20 up to 300, in increments of 20 casualties.

The baseline model was modified to investigate each study aim in turn. Aim one required that for each casualty load, the total on-shelf RBC stock inventory at the TC was increased up to 10 times the standard volume by the addition of one complete RBC inventory level each time (Table 3). For aim two a single inventory restock was programmed to occur during each simulation run, providing 100% of the total initial RBC on-shelf stock when activated. This was programmed to occur once per simulation and at time increments of one hour up to 12 hours from the start.

Statistical analysis was performed in GraphPad Prism version 5.01 (GraphPad Software Inc. San Diego, CA, USA) and Microsoft Excel (Microsoft Corp. Redmond, WA, USA). The simulation software’s use of common random numbers in the model allowed scenario comparison using paired t-tests. Statistical significance was measured at a p-value of <0.05 unless stated. Results are given as means and 95% confidence intervals (CI) or medians with the inter-quartile range (IQR) depending on the data type.
Results

The study involved over 35,000 simulations. The first experimental aim was to establish the effect of varying on-shelf RBC stock hold at a TC prior to an event on the ability to adequately treat increasing numbers of casualties. The effect of this variation on two of the outcome measures is shown in Figure 2. Standard TC stock hold is inadequate in terms of meeting complete casualty treatment for even the smallest event size. This 20 casualty load equated to 12 P1s and 8 P2s with approximately 10 and 4 of these requiring RBCs respectively. 35% (CI 32-38) of bleeding P1s are left inadequately treated with this casualty number (Figure 2A). Bleeding P2s, despite a lower overall RBC demand are also left untreated with 7% (CI 4-10) failing to receive their RBC demand within four hours. In addition, exhaustion of type O emergency RBC stocks (a surrogate for reaching surge capacity) occurred in a median of 10 hours (IQR 5->12) under these conditions (Figure 2B).

Doubling the casualty load to 40 increases the percentage of bleeding casualties left inadequately treated to 60% of P1s (CI 57-63) (Figure 2A). P2s are also affected with a rise 7% to 30% (CI 26-34) left untreated and the time to reach surge capacity is reduced to approximately 2 hours (IQR 1-3) (Figure 2B). Doubling the standard on-shelf RBC hold ensured the complete treatment of all bleeding P2s within four hours with a 20 patient casualty load. Successive on-shelf stock maintains this treatment rate with each additional 20 casualties received up to 180 casualties, where a ten-fold increase in the standard RBC stock level is required. In contrast, the approximate 8% improvement in the treatment rate of bleeding P1s within one hour in the same
scenario is not sufficient to fully treat all bleeding P1s, and further increases in stock hold produce no further improvement (Figure 2A).

Although treatment remains incomplete amongst the P1 cohort irrespective of the RBC inventory, larger stock holds do maintain treatment rates above 50% as casualty load increases (Figure 2A). In order to determine whether given more time than the defined one hour limit, further bleeding P1 casualties would be completely treated, we considered all bleeding casualties treated within six hours of their arrival. This confirmed P1 casualties do go on to receive their full transfusion demand. Meeting bleeding P1 RBC demands is therefore limited by both RBC stock adequacy and the ability to provide this stock within the required hour.

The consumption of emergency type O RBC follows a similar stepwise pattern to the P1 and P2 treatment rates, requiring ever increasing multiples of stock to ensure adequate levels are maintained as the casualty load grows (Figure 2C). Double the standard stock is required to ensure adequacy of supplies following a 20 casualty MCE, thereafter, roughly another 100 units of type O are required with every additional 20 casualties to maintain this state. Without this, a rapid exhaustion of stock occurs within hours of the first casualty arriving across all casualty loads, reiterating the dependence on type O RBCs early in the event response.

In order to plan for future events, transfusion planners need to understand the treatment capacity of the RBC stock they hold. The maximum treatment level of P1s within an hour attainable in the
model was found to be just over 70% irrespective of stock hold. We found 12U of RBC per casualty received (i.e. total casualty load) were required to maintain close to the maximum attainable treatment levels within the priority based time constraints. For an MCE comprising 50 casualties this would equate to 600U of RBCs on the shelf at a TC, far and beyond even the largest hospital based RBC inventories. Accepting a minimum treatment percentage of 90% for all bleeding casualties within six hours of their arrival, the RBC volume required falls to around 9.5U per casualty received (Figure 3). Restocking supplies may therefore be the only option for managing events of any significant size.

Having established the effect on the model outcomes of varying a TC’s RBC stock levels, we wished to explore the second study aim to determine the effect of RBC restocking compared with maintaining equivalent on-shelf stock levels. The maximum RBC stock level achievable in this scenario was twice the standard level held, we therefore focused on casualty loads up to 100 casualties (approximately 48 bleeding P1s and 20 bleeding P2s). Beyond this, the imbalance between stock and casualty load becomes limited in its interpretation (Figure 2). The results for the five casualty loads considered are shown in Figure 4.

P1 casualties treated within an hour showed no statistically significant difference in their outcomes between stock hold and restocking within one hour of the simulation start, up to a casualty load of 80 on paired t-test analysis. However, at a two hour restock time the model displayed a statistically significant reduction in treatment rate once casualty loads reached 60 or greater (Figure 4A). Whilst all casualty loads showed a downward trend in the treatment of
bleeding P1s with increasingly later restocking, the rate of bleeding P2s treated within four hours initially increased above that achieved with an equivalent increased stock hold between casualty loads of 40 and 80. This was maximal with a casualty load of 60, showing a statistically significant 7% increase in the number of bleeding P2s treated when a complete restock was performed at four hours. Beyond four hours there is gradual decline in bleeding P2 treatment rates with increasingly delayed RBC restocking time.

P1 treatment rates showed a gradual decline with increasing restock time, whilst P2s displayed an initial improvement in treatment rates. This was due to the lengthier time window (<4 hours) afforded to P2s to be treated, allowing greater use of group specific RBCs within this cohort, preserving emergency type O supplies. If type O RBCs are available on-shelf they are used within the model, whereas when supplies exhaust, and prior to any further RBC delivery, casualties are forced to wait for group specific blood. This relieves the burden on type O RBC supplies when they are eventually restocked. Despite the negative trend of P1 treatment levels as restock time increases, the overall treatment rate of all bleeding casualties within six hours showed no statistically significant reduction in treatment when a restock occurred within six hours up to a casualty load of 80 (Figure 4B).

**Discussion**

This study set out to investigate strategies for optimizing in-hospital RBC provision to casualties in MCEs using simulation modelling. The study showed standard on-shelf RBC stock levels at TCs to be inadequate for achieving maximum treatment rates possible within the model in even
limited sized MCEs. Holding higher on-shelf stock levels was found to produce more acceptable
treatment levels, however, maintaining the vast volumes of RBC permanently on-shelf is both
expensive and potentially wasteful. There is particular pressure on type O RBC supplies in these
events and the rate of type O consumption closely reflects TC treatment capacity when supplies
cannot be replenished. Restocking of RBC supplies appears to offer a feasible solution within the
model. This is however a time critical process, to avoid a reduction in overall RBC treatment
outcomes the process must be completed within 6 hours of the first casualties’ arrival.

The first aim of the study examined the effect of RBC stock hold volume at the TC on its ability
to meet the RBC demands of arriving casualties. The model reinforced findings from previous
studies that the capacity of responding units to cope with these events maybe more limited than
we realise (39). The standard stock hold of 210U was inadequate for casualty loads as low as 20,
correlating with the RBC consumption of 160U by seven bleeding casualties during the London
2005 bombings which also involved multiple RBC restocks (12). Achieving acceptable levels of
care in the model following such events requires a significant increase in stock availability, most
notably of emergency type O RBCs. When these stocks begin to exhaust, the ability to
effectively manage casualties rapidly deteriorates. The level of this resource and its rate of
consumption is likely to offer the best indication of the transfusion system reaching surge
capacity.

The second aim of this study investigated the effects of restocking supplies during an event in
place of holding the equivalent RBC stock level permanently on-shelf at TCs. The experimental
findings appear to support the use of restocking, with comparable outcomes identified in the model when performed early following an event. Many units have relied on restocking RBCs to manage demand in past MCEs (40). This however, has always been on a request first basis from the TC’s themselves. This approach risks delaying the delivery of RBCs, potentially missing the narrow window of opportunity for improving transfusion related outcomes in these events. The automatic immediate dispatch of RBC supplies from central supplies at the time of an event’s declaration, using a push over pull strategy may ameliorate this risk.

The investigation of aim two illustrated the differing effects on P1 and P2s of exhausting type O RBC stocks for a period of time. The delay in treatment forces greater use of group specific blood and whilst P1 treatment levels deteriorate, P2s, who have a longer treatment window, see an increase in their overall treatment numbers within the model. This suggests more rigid control over access to emergency type O RBCs based on triage priority may offer potential opportunities for preserving RBC stocks in future events.

**Limitations**

Several study limitations should be appreciated when interpreting the results. Firstly, the study was performed on a model and must not be mistaken for real-life experiments. Secondly, adequate treatment was defined as full receipt of RBC demand within set timescales based on triage category. These literature-based definitions in reality refer to time to initiate treatment; however, they do provide a suitable time based target and a quantifiable measure for evaluating system performance. Finally, the provision of components other than RBCs was not considered
in this RBC specific model which could have significant effects on RBC consumption. The rationale for this was that given recent advances in the management of trauma haemorrhage, availability of accurate coagulation therapy data with which to inform the model was highly limited.

**Conclusion**

This study has demonstrated the potential effect adequacy of RBC supplies has on the ability to meet the transfusion demands of bleeding casualties in MCEs through a simulation modelling approach. Whilst greater and greater stock volumes are able to maintain adequate rates of treatment, the logistical and financial implications of maintaining these volumes of RBCs at TCs limits the feasibility of this solution long-term. Restocking RBCs early during an event can produce equivocal outcomes compared to an on-shelf stock hold, however, in order to achieve this, a push over pull approach needs to be considered to prevent delay and maximise any potential benefits.
References

15. Peace TIfEa. The Global Terrorism Index. IEP Reports. 2015;36.


Table 1: Assumptions applied in the model design

<table>
<thead>
<tr>
<th></th>
<th>Assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The TC was in a state of readiness for arrival of the first casualty with no other current casualty encounters in progress.</td>
</tr>
<tr>
<td>2</td>
<td>Casualties were correctly triaged upon arrival and no variation in their triage priority occurred during the simulation.</td>
</tr>
<tr>
<td>3</td>
<td>The triage description definitions of time to treat P1 and P2s were taken as time from their arrival into the model and not the time of the initial event.</td>
</tr>
<tr>
<td>4</td>
<td>Only P1 and P2 casualties required transfusion and all P3 casualties were treated at a separate location to avoid impacting on care of those more seriously injured.</td>
</tr>
<tr>
<td>5</td>
<td>No system failures or human errors occurred throughout the simulation.</td>
</tr>
<tr>
<td>6</td>
<td>All physical treatment resources aside from RBCs and the components involved in RBC provision to casualties were considered adequate for casualty care. This included medical staff.</td>
</tr>
<tr>
<td>7</td>
<td>All bleeding casualties were assumed to have a defined RBC demand from admission which did not increase during their time in the model.</td>
</tr>
<tr>
<td>8</td>
<td>Casualty blood samples did not require collection or delivery as they were transported by automatic pneumatic air tube systems.</td>
</tr>
<tr>
<td>9</td>
<td>Emergency type O RBCs were available for immediate transfusion from satellite blood fridges kept in most acute treatment areas whereas type specific units required collection.</td>
</tr>
</tbody>
</table>
Table 2: A summary of the data inputs for each simulation variable

<table>
<thead>
<tr>
<th>Input Parameter</th>
<th>Input Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Casualty Arrival Rate</td>
<td>Johnson SB Distribution ($\gamma 1, \delta 0.55, \lambda 15, \Upsilon 0.05$)</td>
</tr>
<tr>
<td>Casualty Load</td>
<td>20-300 in 20 casualty increments per run</td>
</tr>
<tr>
<td>Casualty Blood Type</td>
<td>A = 42%, B = 10%, AB = 4%, O = 44%</td>
</tr>
<tr>
<td>Proportion of P1:P2s</td>
<td>Percentage Probability: P1 = 60%, P2 = 40%</td>
</tr>
<tr>
<td>Assessment and Access Time</td>
<td>Gamma Distribution P1 = ($\alpha 2.1, \beta 5.0$), P2 = ($\alpha 1.7, \beta 11.0$)</td>
</tr>
<tr>
<td>Proportion of P1 &amp; P2s Bleeding</td>
<td>Percentage Probability: P1 = 80%, P2 = 50%</td>
</tr>
<tr>
<td>P1 &amp; P2 RBC Demand</td>
<td>Poisson Distribution: P1 = (10.7), P2 = (4.7)</td>
</tr>
<tr>
<td>Blood Sample Transport</td>
<td>Constant 3 minutes</td>
</tr>
<tr>
<td>Book &amp; Verify Sample</td>
<td>Constant 1 minutes</td>
</tr>
<tr>
<td>Centrifuge Sample</td>
<td>Constant 5 minutes</td>
</tr>
<tr>
<td>Verify Sample &amp; Load Analyser</td>
<td>Constant 1 minutes</td>
</tr>
<tr>
<td>Sample Analysis</td>
<td>Constant 11 minutes</td>
</tr>
<tr>
<td>Unload Sample &amp; Verify</td>
<td>Constant 1 minutes</td>
</tr>
<tr>
<td>Dispense Grouped RBC</td>
<td>Constant 30 seconds per Unit of RBC</td>
</tr>
<tr>
<td>Delivery of RBC</td>
<td>Constant 5 minutes</td>
</tr>
<tr>
<td>Transfusion Time</td>
<td>Johnson SB Distribution: P1 = ($\gamma 2.1, \delta 0.75, \lambda 29.0, \Upsilon 0.56$), P2 = ($\gamma 1.2, \delta 0.62, \lambda 29.0, \Upsilon 0.84$)</td>
</tr>
</tbody>
</table>

*Where a probability distribution is provided, a value is sampled from this distribution each time it is referenced within the model. Distribution parameters are provided in brackets.*
Table 3: Volume of each type of blood group held on-shelf at the start of the simulation for each multiple of the basic inventory state (shown as units of RBC)

<table>
<thead>
<tr>
<th>Blood Group</th>
<th>Multiples of Basic Inventory Stock Level</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type O RBC</td>
<td></td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>400</td>
<td>500</td>
<td>600</td>
<td>700</td>
<td>800</td>
<td>900</td>
<td>1000</td>
</tr>
<tr>
<td>Type A RBC</td>
<td></td>
<td>75</td>
<td>150</td>
<td>225</td>
<td>300</td>
<td>375</td>
<td>450</td>
<td>525</td>
<td>600</td>
<td>675</td>
<td>750</td>
</tr>
<tr>
<td>Type B RBC</td>
<td></td>
<td>25</td>
<td>50</td>
<td>75</td>
<td>100</td>
<td>125</td>
<td>150</td>
<td>175</td>
<td>200</td>
<td>225</td>
<td>250</td>
</tr>
<tr>
<td>Type AB RBC</td>
<td></td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>40</td>
<td>50</td>
<td>60</td>
<td>70</td>
<td>80</td>
<td>90</td>
<td>100</td>
</tr>
<tr>
<td>Total RBC Stock</td>
<td></td>
<td>210</td>
<td>320</td>
<td>630</td>
<td>840</td>
<td>1050</td>
<td>1260</td>
<td>1470</td>
<td>1680</td>
<td>1890</td>
<td>2100</td>
</tr>
</tbody>
</table>
LIST OF FIGURE LEGENDS

Figure 1: An Activity Diagram using UML of the model structure.

Figure 2: The relationship between the number of overall casualties received at the TC, the multiples of standard all type RBC stock held on-shelf at the centre and: A) The percentage of bleeding P1 casualties receiving full treatment within one hour, B) The median time to exhaustion of emergency type O RBC supplies. *Indicates stock remained available through to the end of the simulation run (72 Hours).

Figure 3: The relationship between units of RBC held per casualty received and the percentage of all bleeding casualties treated within 6 hours.

Figure 4: Examined through the first five casualty loads modelled of 20, 40, 60, 80 and 100, the effect of restocking a TC’s RBC inventory at increasing hourly time points from the time of first casualty arrival on: A) The percentage of bleeding P1s treated under 1 hour, and B) The percentage of bleeding P1 and P2s treated under 6 hours. Values are shown as means with 95% confidence intervals.
Percentage of Bleeding Pts. Treated in Under 1 Hour

Multiples of Held RBCs Stocks

Total Number of Casualties Received
Figure 2

B
Figure 3

The graph illustrates the percentage of bleeding casualties treated within 6 hours as a function of the units of red blood cells (RBC) per overall casualty received. Key levels are marked at 50%, 75%, 90%, and 100%.

- **100% Level**: Achieved at approximately 25 units of RBC per overall casualty.
- **90% Level**: Reached at about 20 units of RBC per overall casualty.
- **75% Level**: Attained at around 15 units of RBC per overall casualty.
- **50% Level**: Reached at about 10 units of RBC per overall casualty.

The graph shows a steady increase in the percentage of treated casualties as the units of RBC per overall casualty increase.
LEGEND: Casualty Loads: 20 (□), 40 (■), 60 (X), 80 (●) and 100 (○)
Figure 4

LEGEND: Casualty Loads: 20 (□), 40 (■), 60 (X), 80 (●) and 100 (○)