Virtual Competition Rules:

Challenge Event 1

Version 1h

July 9, 2024

Defense Advanced Research Projects Agency
Biological Technologies Office
675 North Randolph Street
Arlington, VA 22203-2114
TriageChallenge@darpa.mil

1 Contents

2 Introduction

This document describes the Challenge Event 1 Rules of the DARPA Triage Challenge (DTC). This document supersedes previous versions of the DARPA Triage Challenge Rules. Significant revisions from past versions in this document are indicated by blue text. Teams are encouraged to closely review the entire document. The intent of this document is to provide participants guidance on competition design and scoring objectives to inform their development efforts in preparation for the first competition event. This document is subject to change and may be superseded by later versions. The latest official versions of all documents are posted on the DARPA Triage Challenge Website (triagechallenge.darpa.mil) and the DARPA Triage Challenge Community Forum.

DARPA intends to release a draft of the Competition Rules no later than nine months before each Challenge Event. The final version of the Competition Rules will be released no later than three months prior to each respective event.

The DARPA Triage Challenge Chief Judge has the final authority to make any decisions related to the rules or scoring. All decisions made by the Chief Judge are final.

The main goal of the DARPA Triage Challenge is to inspire development of scalable, timely, and accurate capture of novel injury signatures to enhance triage decision-making in austere, complex, and mass-casualty settings. The challenge elements and the competition structure itself are intended to address the additional goal of increasing the diversity, versatility, cost-effectiveness, and robustness of relevant technologies and systems capable of addressing the myriad needs of a wide range of mass casualty incidents (MCIs) rather than single-purpose or specifically tailored solutions. The third goal of the DARPA Triage Challenge is to establish a collaborative community by bringing together multi- disciplinary teams and cross-cutting approaches across disparate fields to address the autonomy, perception, and diagnostic needs of the medical triage community.

3 Overview

Under the authority of 10 U.S.C. §4025 to stimulate innovations using prize competition, the DARPA Triage Challenge will use a series of competition events to drive breakthrough innovations in the identification of physiological features ("signatures") of injury. These new signatures will help medical responders perform scalable, timely, and accurate triage. Of particular interest are MCIs, in both civilian and military settings, when medical resources are limited relative to the need.

The DARPA Triage Challenge's long-term vision is 1) an initial, or primary stage of MCI triage supported by sensors on stand-off platforms, such as uncrewed aircraft systems (UASs) or uncrewed ground systems (UGSs), and algorithms that analyze sensor data in real-time to identify casualties for urgent hands-on evaluation by medical personnel; followed by 2) a secondary stage, after the most urgent casualties have been treated, supported by non-invasive sensors placed on casualties and algorithms that analyze sensor data in real-time to predict the need for life-saving interventions (LSIs) by medical personnel. Injury information provided by these sensors in primary and secondary triage could be integrated with other information about the scene to accumulate evidence about the injury mechanism and characteristics in order to enhance overall situational awareness, and to focus further physiological interventions.

To advance progress towards this vision, the DARPA Triage Challenge aims to bring together multi-disciplinary teams and industries that will identify physiological signatures and develop sensor and algorithm strategies for complex MCI settings. Teams participating in the DARPA Triage Challenge will be tasked with developing and demonstrating strategies for capturing high-value signatures for either primary

[1] Patterns in sensor data that reflect or predict injuries of high importance for triage assessments

or secondary triage, or for both. While aspects of the DARPA Triage Challenge involve sensors and sensor-delivery platforms, the priority is the development of physiological signatures and models to detect them, not the development of new sensor or platform technology.

4 DARPA Triage Challenge Schedule Overview

The DARPA Triage Challenge is a 3-year effort with 3 sequential 12-month phases for Primary Triage (Systems and Virtual Competitions) and Secondary Triage (Data Competition) in parallel, each culminating in a challenge event (Figure 1; see the DTC website for competition details). In each phase, competitors will develop signatures and detection and analysis strategies for each Competition. DARPA will host two competition events in each phase; a workshop and a challenge event.
Competition events will become progressively more difficult and realistic from Phase 1 to Phase 3.

The workshops will provide an opportunity for practice runs for all tracks and an opportunity for Systems competition teams to collect data from physical simulations of scenarios similar to the end-of-phase challenge event.

Table 1 provides additional information on schedule and format of Competition events and workshops.

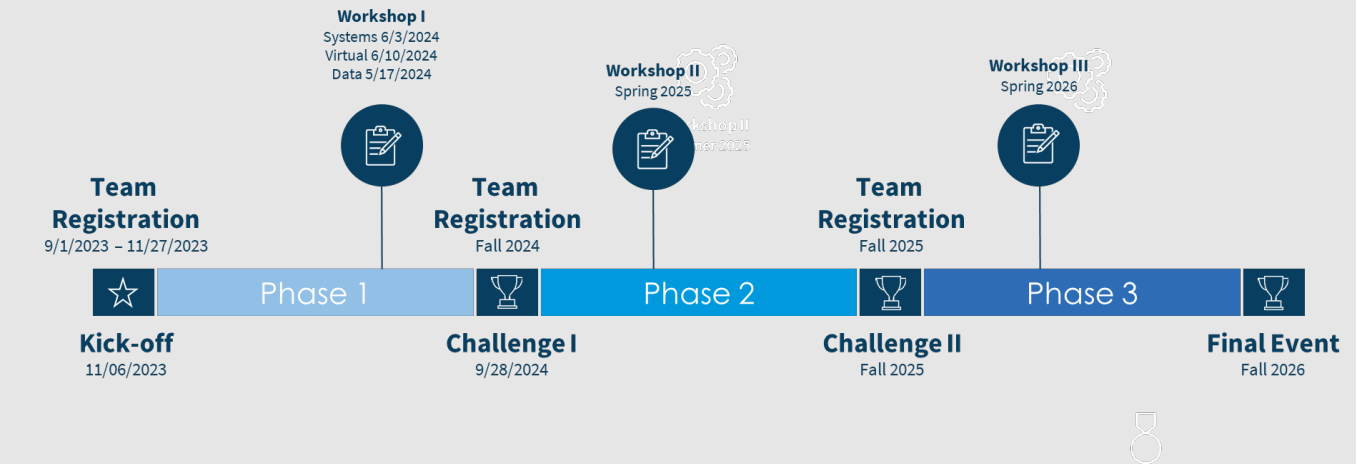

Figure 1- Program structure and schedule for the DTC.

Virtual Competition and Data Competition - Tracks C, D and E			
Event	**Location**	**Est. Duration**	**Date**
Year 1			
Challenge Kick-off	Hybrid	2 days	Nov 6-7, 2023
Workshop - Month 8 *Lessons-learned session*	Virtual	1 day	6/17/2024
Challenge 1 - Month 12 *Evaluations / runs*	Virtual	1 month	7/30/2024-8/30/2024
Challenge 1 - Month 12 *Awards /lessons-learned session*	Hybrid	1 day	10/5/2024
Year 2			
Workshop - Month 4	Hybrid	1 day	Spring 2025
Challenge 2 - Month 12 *Evaluations / runs*	Virtual	TBD	Fall 2025
Challenge 2 - Month 12 *Awards /lessons-learned session*	Hybrid	1 day	Fall 2025
Year 3			
Workshop - Month 4	Hybrid	1 day	Spring 2026
Final Challenge - Month 11 *Preliminary rounds*	Virtual	TBD	Fall 2026
Final Challenge - Month 11 *Finalists only - Runs and Awards*	In person	1 day	Fall 2026

Table 1 - Schedule of DARPA-organized Challenge events and workshops..

5 Prizes and Funding

Teams are encouraged to pursue high-risk, high-reward approaches to meet and exceed the objectives of the Challenge Events. Monetary prizes will be awarded for the Systems, Virtual, and Data Competitions at each of the Challenge Events (Table 2).

Challenge I — Fall 2024

Systems [self-funded]	Virtual [self-funded]	Data [self-funded]
1st $120,000	1st $60,000	1st $120,000
2nd $60,000	2nd $30,000	2nd $60,000
3rd $20,000	3rd $10,000	3rd $20,000

Challenge II — Fall 2025

Systems [self-funded]	Virtual [self-funded]	Data [self-funded]
1st $300,000	1st $300,000	1st $300,000
2nd $150,000	2nd $150,000	2nd $150,000
3rd $50,000	3rd $50,000	3rd $50,000

Challenge III — Fall 2026

Systems [DARPA-Funded and self-funded]	Virtual [Self-funded]	Data [DARPA-Funded and self-funded]
1st $1,500,000	1st $600,000	1st $900,000
2nd $750,000	2nd $300,000	2nd $450,000
3rd $250,000	3rd $100,000	3rd $150,000

Table 2 – Prize structure for the three Challenge Events

DARPA-Funded Teams
DARPA-funded teams (Tracks A and D) are only eligible for the prizes in the Final Events (selection for DARPA-funded team has closed). The Government's obligation for prizes under DTC is subject to the availability of appropriated funds from which payment for prize purposes can be made. No legal liability on the part of the Government for any payment of prizes may arise unless appropriated funds are available to DARPA for such purposes.

Self-Funded Teams
Self-funded teams (Tracks B, C, and E) are eligible for prizes in all of the Challenge Events.

Self-funded Systems Competition Teams Prizes and Funding: The Phase 1 and Phase 2 prizes for self-funded Systems teams will be awarded to the best performing self-funded Systems Teams in each event, provided that the teams finish in the top 5 overall (including DARPA-funded) teams. High-performing self-funded teams are also eligible to become a DARPA-funded Track A team in Phase 2 and/or 3. The Government's obligation for prizes under DARPA Triage Challenge is subject to the availability of appropriated funds from which payment for prize purposes can be made. No legal liability on the part of the Government for any payment of prizes may arise unless appropriated funds are available to DARPA for such purposes.

Virtual Competition Prizes and Funding: The Phase 1 and Phase 2 prizes for the Virtual Competition will be awarded to the best performing teams. The Government's obligation for prizes under DARPA Triage Challenge is subject to the availability of appropriated funds from which payment for prize purposes can be made. No legal liability on the part of the Government for any payment of prizes may arise unless appropriated funds are available to DARPA for such purposes.

Self-funded Data Competition Prizes and Funding: The Phase 1 and Phase 2 prizes for self-

funded Data Competition teams will be awarded to the best performing self-funded Data Teams, provided that the teams finish in the top 5 overall (including DARPA-funded) teams. High-performing self-funded teams are also eligible to become a DARPA-funded Track D team in Phase 2 and/or 3. The Government's obligation for prizes under DARPA Triage Challenge is subject to the availability of appropriated funds from which payment for prize purposes can be made. No legal liability on the part of the Government for any payment of prizes may arise unless appropriated funds are available to DARPA for such purposes.

To be eligible for prizes, teams must first be registered in the team qualification portal. The award process requires recipients to furnish information that may trace or identify recipients either individually or as an organization (e.g., Social Security Number or Tax Identification Number). The primary contact of each registered team is responsible for providing the award information necessary for prize disbursement. DARPA will reach out by email to the primary contact of each registered team to either confirm their vendor status or request the required forms (e.g., SF-3881 or PIF). DARPA is not responsible for disbursement of prizes to any team members other than the primary contact/organization.

At the end of each competition event, teams will be invited to discuss their technical approaches and lessons learned in a townhall-style hotwash. The extent of technical details shared does not need to exceed data agreements established upon qualification.

6 Qualifications

Prospective DTC competitors must demonstrate track-appropriate performance capabilities to be eligible to participate in DARPA Triage Challenge. All teams in all three competitions (Primary Triage Systems tracks, Primary Triage Virtual track, and Secondary Triage Data tracks; see the DTC website for track details) must complete two types of qualification: a Team Qualification at the beginning of each phase, and a later event-specific Event Qualifications for each Workshop and Challenge Event. Successful Team Qualification is a prerequisite to Event Qualifications in the same phase.

The initial *DTC Event Qualification Guide* is expected to be released by February 18th, 2024. The *DTC Event Qualification Guide* will continue to be updated for each event. The latest revision will be posted on the DTC Website and DTC Community Forum.

6.1 Team Qualification

Teams must qualify for DARPA Triage Challenge competition events during the designated qualification window by completing the *Team Qualification* form on the DTC Team Portal. Team Qualification submissions will be accepted on a rolling basis but must be submitted by the deadline (3). Team qualification is required to receive access to datasets and prior to event-specific enrollment.

Team Qualification Windows by Phase	
Phase 1	9/1/2023 - 11/13/2023
Phase 2	9/1/2024 - 11/15/2024
Phase 3	Fall 2025

Table 3 – Team qualification schedule.

6.2 Event Qualification

Prospective teams are required to demonstrate baseline performance and utility capabilities (e.g., safety

measures for the Systems Competition, simulator usage for the Virtual Competition, and algorithm capability for the Data Competition), to be eligible to participate in events. **All** teams (DARPA-funded and self-funded) in all competitions (Systems, Virtual and Data) must qualify for each event including the DTC workshops, Preliminary Events (i.e. Phase 1 and Phase 2 Challenge Events), and Final Event.

The latest revision of the *DTC Event Qualification Guide* will be posted on the DARPA Triage Challenge Website and DTC Discourse Community Forum. Event Qualification submissions will be accepted on a rolling basis but must be submitted by the deadline to be eligible to participate in the event (Table 4). The specific qualification deadlines for each event are provided in the *DTC Event Qualification Guide.*

Failing a previous qualification attempt does not preclude a team from resubmitting a revised qualification submission within the qualification deadlines for any given event. DARPA may adjust the qualification rules for each event and may choose to award qualification waivers for teams that have successfully participated in a prior Workshop or Challenge Event.

DARPA reserves the right to disqualify any team that is found to violate either the rules or applicable laws and regulations.

Event	Event Qualification	Event Date
Workshop 1	3/5/2024 - 4/5/2024	6/3/2024 - 6/8/2024
Challenge 1	6/28/2024 – 7/30/2024	7/30/2024 - 8/3/2024
Workshop 2	12/5/2024 -1/5/2025	March 2025
Challenge 2	Summer 2025	Fall 2025
Workshop 3	Winter 2025-2026	Winter 2025-2026
Challenge 3	Summer 2026	Fall 2026

Table 4 – Event qualification schedule.

7 DARPA Triage Challenge Technical Workshops

DARPA encourages vibrant information exchange and collaborative interactions among all DARPA Triage Challenge participants, to include DARPA technical staff, independent verification and validation (IV&V) teams, representatives from competitor teams, infrastructure developers, and other government partners. To that end, DARPA will host a virtual workshop for the Virtual Competition in each phase which will offer a forum for community building and cross-pollination of technical ideas and approaches.

In each phase (8 months into Phase 1, 4 months into Phases 2 and 3) DARPA will host a multi-day hybrid workshop. This will include live practice sessions for Systems Competition competitors to test their systems on simulated casualty scenes similar to the next challenge event and virtual practice sessions for Virtual Competition teams to test their algorithms in simulated casualty scenes similar to the next challenge event. The practice sessions will be followed by a virtual 'lessons learned' discussion for all tracks and an opportunity to discuss real-world needs with Government partners.

At the workshops, teams will have opportunities to confirm integration with the DARPA virtual platform and scoring systems, and inform their development efforts. Runs at the workshops are not officially scored, but teams are encouraged to operate according to the Competition Rules to prepare for the Challenge events.

We will hold a virtual lessons learned meeting shortly after the workshop for teams to discuss experience gained regarding technical aspects of their systems at the workshop tests.

8 Human Subjects Research (HSR)

For the Virtual Competition, use of training data provided by DARPA does not constitute HSR, and competitors do not need to obtain IRB approval to use these data. **Self-funded teams are prohibited from the collection or use of any other human subject data as part of their involvement in the DARPA Triage Challenge, beyond data and data-collection opportunities provided by DARPA, because DARPA HSR supervision is not feasible for teams not under DARPA contract.** Self-funded teams should carefully consider this limitation and should take this into account in their technical approach, leveraging other strategies as appropriate (*e.g.*, simulations).

DoD Definition of Human Subjects Research (HSR)

The term "human subject" can be applied to research efforts that meet EITHER of the following criteria: A

living individual about whom an investigator (whether professional or student) conducting research:

- Obtains information or biospecimens through intervention or interaction with the individual, and uses, studies, or analyzes the information or biospecimens; or
- Obtains, uses, studies, analyzes, or generates identifiable private information, personally identifiable information, or identifiable biospecimens.

Human Subjects Research involves:

- Activities that include both a systematic investigation designed to develop or contribute to generalizable knowledge and involve a living individual about whom an investigator conducting research obtains information or biospecimens through intervention or interaction with the individual, or identifiable private information, or biospecimens.

8.1 Handling of DARPA-provided data

Primary Triage:

Primary triage datasets are owned by the Army and developed by its Telemedicine & Advanced Technology Research Center (TATRC), and shared with DARPA under appropriate authorities, exclusively for research purposes (including DTC).-The TATRC datasets entrusted to DARPA have been intentionally de-identified to ensure—to the greatest extent practicable—that there is no reasonable basis to believe that the data could be used to trace a specific identity or present a risk of harm to any individual. However, TATRC datasets may incidentally or unintentionally contain sensitive information and images (including facial imagery). Therefore, as previously acknowledged in the DTC Qualification process, competitors agree they will ***not*** attempt to re-identify, share, or re-use Army/TATRC data as provided by DARPA.

9 Primary Triage: Virtual Competition Rules

9.1 Virtual Testbed

DARPA is investing in the development of the DTC Virtual Testbed, illustrated graphically in Figure 4,

comprising of (1) DTC Simulator, an extensible Unreal Engine (UE5)-based simulation environment employing MetaHumans augmented with physiology data; (2) automated testing and assessment tools; and (3) associated software support infrastructure. This suite of simulation tools is intended to support teams in both the Systems and Virtual Competitions as they develop and evaluate their approaches.

An initial version of the DTC Virtual Testbed and associated simulated sensing videos was released in November 2023. This will be followed by regular releases with updates and enhancements (see Table 5 for the initial roadmap). DARPA intends to continue adding significant improvements and new capabilities to the Testbed over the life of the challenge. Proposed capabilities for the Phase 1 challenge include support for user-generated paths, at least four sensing modalities (RGB camera with audio, thermal camera with audio, LiDAR), and at least 30 simultaneous casualties consisting of at least 4 injury classes (hemorrhage, burn, respiratory distress, and TBI) and at least 20 different human 3D models. The human models will vary by age (16 and older), size (varying BMI), and skin tone (using the 6 types from the Fitzpatrick scale).

Version	Release Date	Changes From Previous Version				Expected Use
		Nav Path	Injury Types	Sensing Modalities	# UGV / UAV Options	
(Phase 0)	11/23	Fixed	Hemorrhage, Burn, Respiratory Distress	RGB Thermal Audio	1 UGS 1 UAS	Phase 1 training / dev
0.1 (Phase 1) – Minor rel	3/24	Manual teleportation			3 UGS 3 UAS	Phase 1 training / dev Phase 1 Workshop Qualification
0.2 (Phase 1) – Workshop	6/24	User Defined Waypoints	Traumatic Brain Injury	LiDAR		Phase 1 Workshop Phase 1 Challenge Qualification
0.9 (Phase 1) – Pre-Challenge	7/24			Audio		Phase 1 Challenge Preparation
1.0 (Phase 1) – Challenge	9/24 (DARPA)					Phase 1 Challenge
	11/24 (all)					Phase 2 training / dev Phase 2 Workshop Qualification
1.3 (Phase 2) – Workshop	3/25	Autonomy Interface	Musculoskeletal Injuries	RADAR Chem Detector	4 UAS	Phase 2 Workshop
1.5 (Phase 2) – Minor rel	6/25		Chemical exposure		4 UGS	Phase 2 training / dev Phase 2 Challenge Qualification
2.0 (Phase 2) – Challenge	9/25					Phase 2 Challenge
	11/25					Phase 3 training / dev Phase 2 Workshop Qualification
2.3 (Phase 3) – Workshop	3/26		Polytrauma	Feedback Solicited	5 UAS	Phase 3 Workshop
2.5 (Phase 3) – Minor rel	6/26				5 UGS	Phase 3 training / dev Phase 3 Challenge Qualification
3.0 (Phase 3) Challenge	9/26					Phase 3 Challenge
	11/26					Future Development

Table 5 – DTC Virtual Testbed Anticipated Release Plan and Capability Growth

The DTC Virtual Testbed will use data from a physiology engine combined with injury timelines to simulate realistic MCIs and UxSs to provide a platform for training primary triage platforms and algorithms (Figure 4). The Testbed will also be used for challenge scoring at the end of each competition. A Virtual Testbed ICD that defines how teams provide input to the testbed (e.g., navigation waypoints / path, platform configuration, sensor configurations, etc.) and the required format / content of the casualty reports / logs was published in early 2024. Any interface changes will be documented via updated versions of the ICD when new releases of the Testbed are provided.

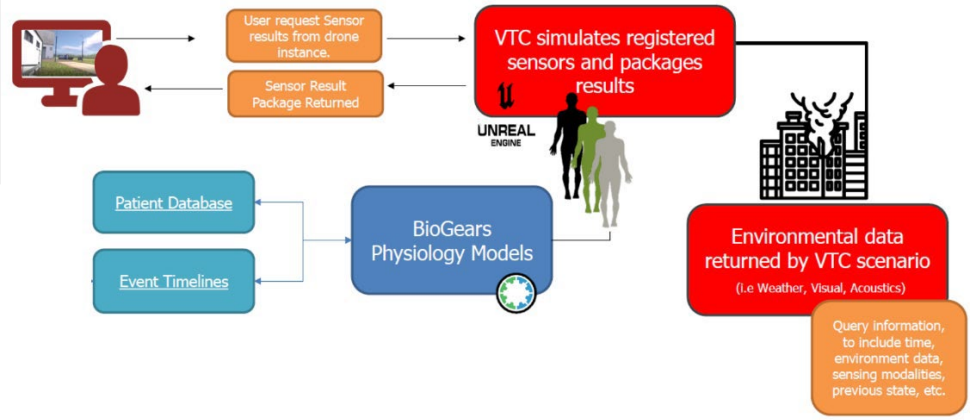

Figure 2- Virtual Testbed

9.2 Virtual - Technical Challenge Elements

The challenge elements used in the Virtual Competition will correspond to those of the Systems Competition to the fullest extent possible.

The Challenge competition courses will be designed to assess performance across various challenge elements, including: *degraded sensing, obscuring obstacles, terrain obstacles, dynamic obstacles, and dynamic casualties*. The challenge elements are expected to become progressively more difficult from Phase 1 to Phase 3.

1. *Degraded Sensing*: The courses are expected to include elements that range from constrained passages to large open fields, lighted areas to complete darkness, and clear to dusty conditions. Sensors will need to have the dynamic range to reliably operate in these environments. Dust, fog, smoke, talking, flashing light, heat spots, and loud background audio effects are within scope of this challenge element. Extreme temperatures, fire, and hazardous materials are not expected to be within scope.

2. *Obscuring Obstacles:* Casualties may be fully visible to partially obscured to completely obscured, such as buried under a shallow layer of rubble. Sensor modalities capable of penetrating rubble will have an advantage in such situations. Stand-off sensor access to skin may be possible but cannot be assumed. Casualties may also be grouped with limbs, torso, or other body parts overlapping.

3. *Terrain Obstacles*: Systems will be required to demonstrate robustness in navigating a range of terrain features and obstacles. Terrain elements and obstacles may include constrained passages, large drops/climbs, inclines, and rubble. UAVs may be subject to atmospheric turbulence. The environments may include natural or human-made materials; structured or unstructured clutter; and intact or collapsed structures and debris.

4. *Dynamic Obstacles*: Responders, "walking wounded", or other physical changes to the environment will test the agility of the system autonomy to identify, reidentify, and assess casualties.

5. *Dynamic Casualties*: Some treatable injuries may rapidly be fatal, so taking too long to find and assess casualties may result in the window for effective LSI to be missed. While competitors are not expected to re-evaluate casualties for changes in status, casualties who are not evaluated within an appropriate timescale may have a change in status (for example, progression of untreated hemorrhage or airway injury).

The scale and complexity of competition courses is expected to increase each year. The design of the first Challenge Event courses is intended to assess the ability of teams to address the variety of challenging environments presented in a post-battle environment. In Year 1, platforms will travel through an area based on user defined waypoints and associated travel times / dwell times. The primary technical challenge element in Year 1 will be dynamic casualties.

9.3 Virtual - Scored Event Submissions

9.3.1 Versions and releases

Applied Research Associates, Inc. (ARA), the DARPA performer providing the Virtual Testbed, will release four major versions; version 0.1 was released in November 2023 (Phase 0), version 1.0 will be released in Fall 2024 (Phase 1), version 2.0 will be released in Fall 2025 (Phase 2), and version 3.0 will be released in Fall 2026 (Phase 3). ARA will also release at least one version between each of these versions. These interim versions will be released prior to the virtual workshops for each challenge year.

The virtual competition scoring runs are expected to be performed by the IV&V teams (i.e., NOT self-evaluated). The challenge teams will not be permitted to witness the evaluations live.

9.3.2 Scored Competition Scenarios

Multiple scenarios will be devised for each simulated environment (e.g., all scenarios are the same incident and same virtual location, but will vary by the location of casualties, human model assigned to a casualty, and injury assigned to each human model).

For Phase 1, there will be up to two simulated environments. Each simulated environment is expected to have at least 3 different casualty scenarios. Following the completion of scoring, the run scores, and logs will be released. Following the completion of phase 1, the event scenarios will be made available to competitors for training.

Each qualified team must submit a single solution to be scored. The submitted solutions will be evaluated against $\boldsymbol{m}$ number of competition scenarios to test the versatility of the solutions. Each competition scenario will, in turn, be evaluated over $\boldsymbol{n}$ replications (reps) to account for random variability. See Figure 3. In Phase 1, $\boldsymbol{m} = 6$ and $\boldsymbol{n} = 1$.

Competitors will be scored based on their ability to correctly evaluate casualties in the simulation. Casualties will be presented at different severities with both physical indicators of injury and severity (i.e.

lacerations and burns) as well as underlying physiological manifestations of injury (i.e. changes in respiration and heart rate). Casualties will be scannable in multiple sensing modalities, beginning with RGB, thermal, and audio at a minimum for Phase 1.

The Event Score of the ***m*** x ***n*** runs is given by:

$$Event\ Score = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{\sum_{j=1}^{n} run\ score_{ij}}{n}\right)$$

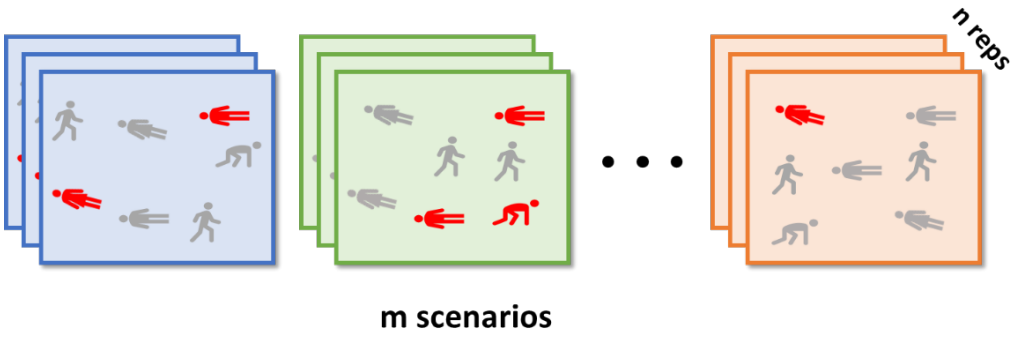

Figure 3 - Virtual Competition approach to scoring competition scenarios

Competitors will be scored based on their ability to correctly evaluate casualties in simulation. Casualties will be presented at different severities with both physical indicators of injury and severity (i.e. lacerations and burns) as well as underlying physiological manifestations of injury (i.e. changes in respiration and heart rate). Casualties will be scannable in multiple sensing modalities, beginning with RGB, thermal, and audio at a minimum for Phase 1.

9.3.3 Solutions Submissions

For scored event submissions, it is expected that qualified teams will submit a single solution in the form of a Docker Image to Amazon's Elastic Container Registry, from where they will be pulled and evaluated against the competition scenarios. ARA will provide a Nextcloud folder for manual uploads as a failsafe. Credentials and specific details will be provided to teams either on the Discourse forum or directly via email. Submissions must be self-contained and will be evaluated through an automated process. Entries that require additional user input or external commands will not be scored. See the Virtual ICD for further details. For Phase 1, competitors must also submit waypoint lists and dwell time for each environment in the competition (see 9.5.1).

The solution submission window for the first challenge will open approximately 2 months prior to the awards ceremony. The submissions will be evaluated and the final results will be announced alongside the Systems Competition results in Fall 2024.

Challenge Event	Submission Window	Results Release
Challenge 1	*7/30/2023-8/30/2024*	*10/5/2024*

Challenge 2	*Summer 2025*	*Fall 2025*
Challenge 3	*Summer 2026*	*Fall 2026*

Table 6 – Submission window for the Virtual track competitors

9.3.4 Human Supervisor

The submitted solutions will be evaluated with no external operator interfaces such as command line inputs or user interventions. Virtual Teams are required to develop self-contained solutions that navigate, search, and evaluate entirely autonomously without human supervisor interactions. For Challenge 1, navigation will be handled via competitor-selected waypoints as described in 9.5.1 Navigation.

9.3.5 Staging Area

Each scored scenario begins with a Staging Area similar to the Systems Competition (see Systems Competition Rules), inside which all platforms in the team may spawn. Platforms will not be spawned outside of the staging area. At spawn time, platforms are provided their ground truth location and orientation. Orientation along with a local coordinate system will be used for casualty reporting and navigation.

9.3.6 DTC Simulator Mechanics

The DTC Simulator utilizes a number of simulator mechanics to address the environmental accuracy of the competition and reduce the operational disparity between the Systems and Virtual Competitions. To that end, team submissions must fully utilize and not seek to circumvent use of simulator mechanics.

9.3.7 Run Termination

A scored run terminates upon any of the following conditions:

- Time Expiration: The given time expires before another termination criterion is met
- Completed run: All waypoints have been visited for the full dwell time
- Violation of minimum standoff distance. The user defined path for a robot breaches the minimum standoff distance based on total robot weight of the robot and all attached payload.
    - Less than or equal 250g: 1 meter.
    - More than 250g and less than or equal 5kg: 3 meters.
    - More 5 kg: 5 meters
- Unexpected termination of competitor container

9.3.8 Score Disputes

Score Disputes are intended to provide teams a mechanism to submit a formal dispute or request for review by the Chief Judge. All score disputes must be sent by email to the DARPA Triage Challenge email address (triagechallenge@darpa.mil) within 24 hours of receiving competition log files. All disputes or requests will be reviewed by the Chief Judge in a timely manner. All decisions made by the Chief Judge are final.

9.4 Virtual - Preliminary Event Scenarios

9.4.1 Challenge Event Competition Environments

The scale and complexity of the environments is expected to vary across competition scenarios and across events. DARPA intends to release practice scenarios (in workshops) in advance of the Challenge Events to provide representative environments in which to develop and evaluate solutions.

In the year 1 competition events it is not expected that there will be space constrained portions of the course. In future years, it is possible that some portions of the environments will only be accessible via passages that are approximately one meter in height and/or one meter in width. The constrained spaces will be large enough for the available platforms to traverse but will limit visibility. The most constrained portions of the competition environments are not expected to be immediately at the entrance.

9.4.2 Preliminary Event Casualties

DARPA will announce the anticipated casualty types in advance of each competition event. The casualty types used in the Virtual Competition will be similar to those used in the Systems Competition. Casualty placement throughout the competition scenarios will be consistent across all repetitions for any given scenario but will vary across scenarios.

9.4.3 Challenge Event Run Duration

Each run will be between 15 and 30 minutes of in-simulation time.

9.4.4 Team Configuration

For Phase 1, the team configuration will be limited to one robot platform. In subsequent phases, the Team Configuration will be limited by the same constraints as the Systems Competition (i.e., maximum of 5 deployed autonomous systems). The DTC Virtual Testbed is expected to include a repository of Robot Operating System (ROS)-available mobile robot models and sensor payload models that will be available for teams to compose their team configuration.

9.5 Virtual - Communications and Score Reporting

9.5.1 Navigation

For Phase 1, while autonomy is still being integrated, competitors will specify system movement via waypoint selection. At the start of the competition submission window, competitors will be provided top-down views of the virtual environment(s) applicable for the competition. The top-down map(s) will contain the staging area where drone(s) will start as well as numbered locations indicating the location of each virtual casualty. Along with their Docker image, competitors shall submit an ordered list of waypoints specifying their desired drone path through each map. The exact specification of the waypoint submission is available in the Virtual Track ICD.

After Phase 1, the Virtual Testbed will support full autonomy of unmanned systems via controls exposed through the Robot Operating System version 2 (ROS2). Once autonomy is fully supported, competitors are

expected to fully control their system(s) within a run of a virtual scenario.

9.5.2 Reporting casualties

Similar to the Systems Competition, teams will need to present relevant information (e.g., casualty health state report, casualty location, corresponding timestamps, etc.) to a virtual Base Station to provide near-real-time situational awareness updates and reports that are scored in a similar manner as the Systems Competition. To report a casualty for scoring, the casualty report must originate from a platform, be sent to the virtual Base Station, and include the location of the casualty being reported. Any casualty reports that do not include a location will be ignored and receive a score of 0.

Casualty reports must identify the location of the casualty to be reported on, the simulation time associated with the report, a competitor-determined casualty ID, and a health assessment. Health assessments do not need to be full and complete; they only need to contain a value for at least one row from Table 7, i.e. a scorable field. For instance, competitors may report heart rate early and subsequently report on injuries and traumas, using the competitor-determined casualty ID to link the separate reports together. Specific details on report format and required fields are found in the ICD.

When a valid report is received for a casualty ID, each scorable field is associated with the report's location and the receipt time of the report. A casualty ID may be reported on multiple times, but, for each scorable field, only the most recent report's value for that field is used for scoring. Said differently, re-reporting a scorable field for a casualty will overwrite any previous reports for that field. At the end of a run, all submitted reports are aggregated and tallied for a final score (see 9.6.2).

The number of competitor-determined casualty IDs tracked by the scorekeeper will be limited to the number of casualties in a scenario. For Phase 1, that number will be 30. Any reports for casualty IDs other than the first 30 will be ignored.

9.5.3 Log Files

At the termination of a run, relevant log files are generated. The log files include all casualty reports (including the casualty's location), their corresponding timestamps (time sensing started, time sending ended, time assessment calculated), score updates, and other details. Additionally, log files allow replaying and viewing of the run by Competition Staff to ensure fair and consistent team performance in the virtual scenario in keeping with the rules and spirit of the DTC.

9.6 Virtual - Scoring Criteria

The goal of the DARPA Triage Challenge is to develop scalable, timely and accurate capture of novel injury signatures to enhance triage decision-making in austere, complex, and mass-casualty settings.

In the Primary Triage challenge (both Systems and Virtual), teams are evaluated based on accuracy and speed in assessing casualty condition using one or more autonomous platforms and stand-off sensors. Casualties are distributed throughout the competition course in a manner which rewards teams that are able to rapidly find and assess casualties. Except for in Phase 1, the location and nature of the casualties is not known in advance of a run by competitors and may be varied from run to run. Results for the Virtual competition will be announced at the prize ceremony on the last day of the competition event. Teams are encouraged to participate in-person for the award ceremonies at Challenge events.

Upon identifying a casualty status, the deployed system must report its injury diagnosis and relevant

clinical information to the evaluation system for scoring. In the Virtual competition, reports will be submitted to the virtual evaluation system. The detailed report format, protocol, and example implementation will be specified in the Virtual ICD. In both competitions, the reported casualty conditions are compared against coincident ground truth data for scoring.

9.6.1 Casualty Localization and Identification

Scoring depends heavily on accurate casualty localization by design. Poor localization may result in the scorable fields for a casualty getting mapped to the wrong ground truth. A ground truth for a casualty is defined as all of the casualty features. To emphasize localization accuracy, submitted reports must be within 5 meters of a ground truth casualty in order to receive any points. In addition, the scorekeeper only allows a ground truth to be associated with a single casualty ID,. A system may submit report updates that overwrite previous reports, so only the final submission for each casualty field will be scored. The golden window bonus points will be based on the time of the last submission of a casualty field.

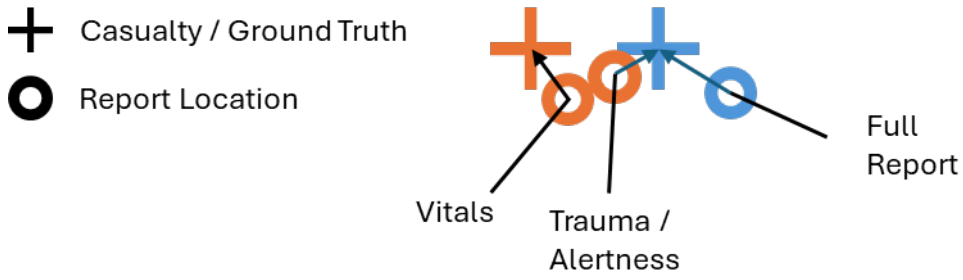

Figure 4. Reports for a single casualty span multiple ground truths. In this case, one report for casualty Orange is closest to the ground truth for casualty Blue. As a result, the ground truth for Blue is "used up" by Orange.

A ground truth is considered "used" when any casualty ID is associated from a report containing one or more casualty report field, regardless of whetherthe report received points. Because casualties will have different physiology, it is disadvantageous to inaccurately localize. Realistically, inaccurate localization will result in receiving few or no points due to matching an incorrect ground truth, and "using up" the value of that ground truth, preventing its use for the casualty it truly represents.

For instance, in Figure 4, assume that the orange casualty is the first one reported on and assume perfect health assessment but imperfect localization. The scorekeeper will award points for vitals but won't award any points for trauma or alertness because the blue ground truth to which they were mapped won't match. Additionally, because the orange trauma and alertness fields are now associated with the blue casualty ground truth, the later full report on blue will receive no credit. In this case, the orange casualty ID is associated to both the orange and blue ground truth casualties.

If a later submission provided an updated location for the orange ID that correctly mapped the trauma and alertness to the orange ground truth it would update the association and leave the blue casualty open for full scoring according to the submitted report with the blue casualty ID.

In the reverse order, if the blue casualty is encountered first, the blue report will receive full credit. The orange vitals report will receive full credit, but the orange trauma and alertness (associated to the blue ground truth casualty) will receive no credit because the blue ground truth was already associated with the blue casualty ID.

In the example below in Figure 5., presume casualties were reported in left-to-right order. The system consistently misjudges the casualties as being closer than they truly are. As a result, the orange casualty's reported location ends up being closer to the green ground truth. The scoring algorithm will map the orange report to the green ground truth. Next, the green report is correctly matched to the green ground truth, but because the orange report was already mistakenly mapped to it, the green report receives no credit. Since casualties won't share equivalent physiology, the orange report will not only destroy the value of the green report by "using" the green ground truth, but also likely receive little to no points of its own due to being mapped to the wrong ground truth. As a result, *both* reports end up receiving little to no points due to inaccurate localization.

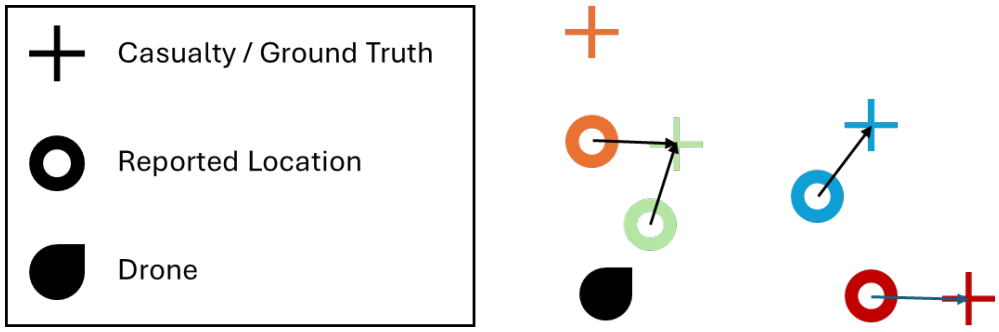

Figure 5. Example scenario where ground truths are incorrectly matched, resulting in devalued scores. In this case, inaccurate localization of orange results in a diminished value of green.

9.6.2 Casualty Reports

Reports capture information about clinical condition of casualties relevant toward triage decision making. The casualty report is comprised of two parts: the first part contains vital signs and time-critical indicators of immediate need, and the second part contains more detailed assessment of injury condition and severity.

In future years, reports may include additional clinical assessment features. The preliminary list of clinical features in the report is shared with the Systems competition and shown in Table 7*, along with definitions provided in* Table 8*. Complete details on report contents and format will be found in the ICD.*

9.6.3 Report Scoring

A valid casualty report earns the team up to 10 points, with up to 5 additional bonus points for early reporting of vitals and time-critical information. Points are determined by comparing submitted report contents to ground truth data from the identified casualty. Categorical fields in the report are awarded points based on whether they match ground truth; numerical fields in the report (i.e., vitals) are awarded points based on whether they are within a predetermined range of the ground truth value. To receive points for Heart Rate, the reported estimate must be within +/- 5 BPM (beats per minute) of ground truth estimated over the preceding 10 seconds. To receive points for Respiratory Rate, the reported estimate must be within +/- 3 BrPM (breaths per minute) of ground truth estimated over the preceding 60 seconds. Physiology is dynamic and can change throughout a simulation run, so all reports must indicate the simulation time they should be compared against in the ground truth. Details on formatting reports, including specifying times, can be found in the ICD.

To incentivize rapid assessment of time-critical information indicating immediate need of medical care, bonus points may be awarded for early casualty reports containing the following fields: Severe Hemorrhage, Respiratory Distress, and vital signs (Heart Rate and Respiratory Rate). Bonus points will be awarded for valid and correct reports received by the DARPA Command Post within an initial "golden

window" during a scored run. The duration of the "golden window" will be half the duration of the run. Teams may receive 2 bonus points each for correct assessment of Severe Hemorrhage and Respiratory Distress, and 1 bonus point for correct assessment of both Heart Rate and Respiratory Rate within the golden window. Note that bonus points will be awarded based on the time the report is received by the virtual base station.

Preliminary scoring criteria and bonus potential for clinical assessment in the casualty report are shown in Table 7. Definitions of casualty report terms are shown in Table 8 and examples can be found in the provided training data. Details about casualty report format and protocol can be found in the ICD.

Field	Values	Scoring Criteria
Severe Hemorrhage[1]	[Present, Absent]	2 if match ground truth (GT) 0 otherwise
Respiratory Distress[1]	[Present, Absent]	2 if match GT 0 otherwise
Heart Rate[2]	Integer	1 if within *n* of GT 0 otherwise
Respiratory Rate[2]	Integer	1 if within *m* of GT 0 otherwise
Trauma	**Head:** [Wound, Normal] **Torso:** [Wound, Normal] **Upper Ext.:** [Wound, Amputation, Normal] **Lower Ext.:** [Wound, Amputation, Normal]	2 if all match GT 1 if at least two match GT 0 otherwise
Alertness	**Ocular:** [Open, Closed, Not Testable (NT)] **Verbal:** [Normal, Abnormal, Absent, NT] **Motor:** [Normal, Abnormal, Absent, NT]	2 if all match GT 1 if at least two match GT 0 otherwise

Table 7 – Preliminary casualty report clinical assessment with scoring criteria

[1] Response receives +2 bonus points if correctly reported within "golden window".
[2] Vitals responses receive +1 bonus point if both are correctly reported within "golden window".

Term	Definition and Indicators in Simulation
Heart Rate	Visual/thermal signals
Respiratory Rate	Chest-wall movement; Visual/thermal signals
Severe Hemorrhage	Indications of severe hemorrhage beyond presence of injury. Any of the following: active bleeding external to the body (e.g., oozing, squirting, pooling), >50% body with blood present on clothes or exposed skin
Respiratory Distress	Indications of respiratory distress beyond Respiratory Rate. Any of the following: tripod position or abnormal head/neck position indicating distress, unequal chest-wall movement, arrhythmic chest movement, gasping sounds with open mouth NOTE: indicated as *Absent* if respiratory rate is 0
Head	Body region including neck and head
Torso	Body region bounded by hips, shoulders, and below neck

Upper Ext.	Body region including arms below shoulders
Lower Ext.	Body region including legs below hips
Trauma – Wound	Any of the following: non-amputation wound (burn, hemorrhage, abrasion, laceration), blood-soaked clothing, hands pressing on wound site
Trauma – Amputation	Traumatic removal of body part with severe hemorrhage at/around wound site
Trauma – Normal	Absence of visual indicators of any wound
Ocular – Open	Eyelids continuously open Eyelids blinking and moving
Ocular – Closed	Eyelids closed and not moving
Ocular – Not testable	Cannot assess due to injured or occluded eyes
Verbal – Normal	Responsive to speech prompts with coherent speech
Verbal – Abnormal	Any of the following: Non-speech or incoherent vocalization, non-responsive to speech prompts
Verbal – Absent	No vocalization
Verbal – Not Testable	Cannot assess due to injured mouth, jaw, and/or throat
Motor – Normal	Purposeful movement of limbs, obeys commands, walking
Motor – Abnormal	Minimal movement or twitching of limbs
Motor – Absent	No movement
Motor – Not Testable	Cannot assess due to sedation or external immobilization

Table 8 – Definitions for fields in the casualty report.

To earn points for correct readings, tolerance for accuracy are specified in table 8.

Vital sign	**Collar**
Heart rate	Rolling averaged over 10 seconds; accuracy of ±5 bpm
Respiration rate	Rolling averaged over 60 seconds; accuracy of ±3 breaths/min

Table 9 – Accuracy collars for heart rate and respiration rate.

As discussed in 9.5.2 and 9.6.1, casualty reports are grouped by competitor-determined casualty IDs and are broken down into the individual scorable fields, which are all stored in the order they are received. A scorable field is essentially a data structure containing:

- Assessment time (to find correct time to compare to within the ground truth)
- Associated ground truth (based on the location of the report this field came in)
- Field ID (e.g. heart rate or Upper Extremity Trauma)
- Field Value (as appropriate for the field)

Upon completion of the run, for each casualty ID, all of the reports and scorable fields for that casualty are aggregated into a single aggregate report that is sent for scoring. The aggregate report is simply the result of taking the latest scorable field for each Field ID. Scoring is summarized with the following pseudo-code:

```
score = 0
for casualty in tracked_casualties
    aggregate = aggregate_reports(casualty)
    for field in aggregate.scorable_fields
        if already_used(field.ground_truth)
            continue  # no points for ground truth re-use
        field_score = get_field_score(field.id, field.value field.ground_truth)
        if in_golden_window(field.receipt_time):
            field_score += get_field_golden_bonus(field.id)
        score += field_score
     mark_ground_truths_used(aggregate)
```

The actual scorekeeper code is available in each of the Virtual Testbed releases.

9.6.4 Time

Time is measured from the start of the run and scoring will be based off of two timestamps: the assessment timestamp and the report timestamp. The assessment timestamp is provided by the deployed system in the report and indicates the time at which the casualty's condition was assessed. This timestamp is used to determine the appropriate ground truth for scoring. The report timestamp is the time when the valid report is received at the virtual base station, and it determines any bonus points awarded.

9.6.5 Final Ranking

For the Primary Triage Virtual Competition, the final ranking will be determined based on each team's event score as described in Section 9.6.3. In the event that multiple teams have an identical event score, those teams will be ranked in ascending order according to the report timestamp of the latest non-zero scoring casualty report.

10 Appendix 1 DTC Glossary

Chief Official – Program manager or higher DARPA authority for the DARPA Triage Challenge.

Systems Competition –Primary Triage Competition run with actors on a real course (Track A, B).

Virtual Competition – Primary Triage Competition run on a virtual platform (Track C).

Data Competition – Secondary Triage Competition (Track D, E).

Chief Judge – DARPA-designated individual with the sole and final authority to make any decisions related to the rules or scoring.

Competition Course – Physical or virtual environment in which deployed systems are expected to explore, and search for casualties.

Judge – DARPA-designated individual with authority to make decisions related to rules, scoring, and safety, with decision-making authority only superseded by the Chief Judge.

Staging Area – Specified area immediately outside of the Competition Course entrance from which teams deploy their system.

Starting Gate – Installed structure or existing entrance which serves as the threshold between the Staging Area and the Competition Course.

Team Lead – Team-designated individual responsible for making team official decisions and communicating with the DARPA Competition Staff.