Predicting Real-World Thoracic Injury Using THOR and Hybrid III Crash Tests

Matthew L. Brumbelow, Jessica S. Jermakian, Raul A. Arbelaez

Abstract In frontal crashes, drivers restrained by a seat belt and airbag are at elevated risk of serious thoracic injury compared with most other body regions. Hybrid III (HIII) sternum deflection has some ability to predict injury, but its utility is limited by several considerations. The Test device for Human Occupant Restraint (THOR) was developed to address limitations of HIII, but it has not been sufficiently validated under combined loading from a belt and airbag. Thirty-five crash tests were conducted with a THOR in the driver seat. Logistic regression was used to assess the ability of THOR metrics to predict injury outcomes in 57 real-world crashes involving matched vehicle designs. Results showed R_{max} was inversely related to Abbreviated Injury Scale (AIS) ≥3 injury outcome, with a 4 mm increase associated with an injury odds reduction of 48% ($p = 0.04$). By contrast, increasing shoulder-belt load was estimated to increase the odds of both AIS≥2 and AIS≥3 injury, with both effects significant at alpha = 0.05. Additionally, several THOR metrics suggesting greater airbag loading were associated with higher R_{max} values in the test data but reduced field injury risk. The biofidelity of THOR under combined restraint loading should be further investigated.

Keywords CISS, Hybrid III, NASS-CDS, THOR, thoracic injury.

I. Introduction

Frontal non-rollover crashes accounted for 50% of fatalities of belted passenger-vehicle occupant in 2019 [1]. This proportion is highest for the newest vehicles (Fig. 1), suggesting that front crashworthiness and/or crash avoidance improvements have lagged behind those for other crash modes. Other than the lower extremities, the thorax is the body region most commonly injured at a level of 3 or greater on the Abbreviated Injury Scale (AIS) [2]. A 2021 study of frontal crashes with an airbag deployment found that a good rating in the Insurance Institute for Highway Safety (IIHS) moderate overlap crash test was estimated to reduce the risk of driver injury for every analyzed body region except the upper extremities and thorax [3]. A separate analysis of field crash data indicated that older drivers of Good-rated vehicles are at high risk of AIS≥3 thoracic injury in crashes with severities similar to the crash test [4]. A 60-year-old driver restrained by a belt and airbag faces an estimated 38% risk of AIS≥3 injury in a crash with a delta V of 70 km/h. The estimated risk rises to 60% when considering all drivers 60yo or older.

The high levels of real-world injury risk are not predicted by Hybrid III (HIII) measurements taken in the IIHS moderate overlap test, also known as the offset deformable barrier (ODB) test. The field injury estimates are based on analysis of over 900 frontal crashes in the National Automotive Sampling System-Crashworthiness Data System (NASS-CDS). These cases are represented by more than 200 different IIHS ODB tests, which produced an average HIII sternum deflection of 31 mm. Using published injury risk curves, this deflection translates to an AIS≥3 injury risk of 12–20% [5-6] (calculated for a 60yo). As these curves correspond to shoulder-belt loading of HIII, while all test and field crashes included airbag loading, the true degree of under-prediction is even greater. If this were simply a scaling discrepancy, it could be handled by lowering the injury rating thresholds in crash tests. However, a subsequent detailed analysis of the NASS cases suggested there may be more fundamental problems with the sternum deflection metric [7] and found that shoulder-belt force, vehicle

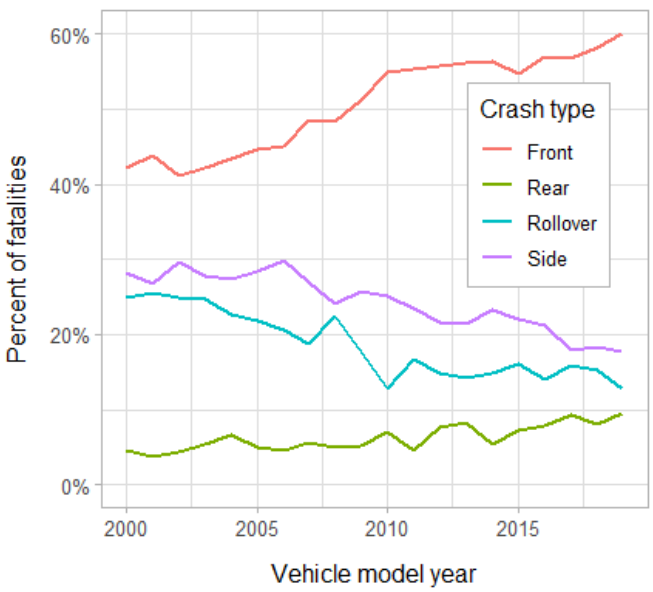

Fig. 1. 2018–2019 US fatalities of belted passenger-vehicle occupants by model year and crash type.

M.L. Brumbelow (mbrumbelow@iihs.org; +1-434-985-4600) is a Research Engineer, J.S. Jermakian is VP of Vehicle Research, and R.A. Arbelaez is VP of the Vehicle Research Center, all at the Insurance Institute for Highway Safety in Ruckersville, VA, USA.

bumper-to-firewall distance, or the ratio between sternum deflection and thoracic acceleration often performed better in predicting injury outcomes than sternum deflection alone.

Limitations of the HIII sternum deflection metric are well documented. First, measured values are highly sensitive to belt placement [8-10]. This makes it problematic to simply reduce rating thresholds when manufacturers can influence belt position, either because they conduct some tests themselves (IIHS verification program) or because they can influence test position of the seat track (the midtrack position in New Car Assessment Program [NCAP] tests) and/or D-ring (both IIHS and NCAP). Second, the single-point measurement often does not represent the response at other locations on the thorax [11-12] and the rotary potentiometer only captures longitudinal deflection. Third, the overall response of the thorax is much stiffer than that of post-mortem human subjects (PMHS) in paired testing [13-14]. For these and other reasons, the risk associated with a given deflection is dependent on the restraint condition [5][8][11], which limits the metric's functionality as a tool for assessing the variety of restraint systems in the vehicle fleet.

The Test Device for Human Occupant Restraint (THOR) Anthropomorphic Test Device (ATD) was developed to address many of the limitations of HIII in assessing risk of injury to the thorax as well as to other body regions. It includes four thoracic deflection measurements taken at different locations using Infra-Red Telescoping Rods for the Assessment of Chest Compression (IR-TRACC), which permit resolution of each deflection into its three-dimensional components. The overall force-displacement response more closely matches PMHS in pendulum impacts [15]. Many other features of the THOR ATD, such as its articulating shoulder, spine stiffness, posture and femur length, were designed to provide a more biofidelic thoracic response to restraint loading in frontal crashes.

Despite THOR's design criteria, its ability to improve thoracic injury prediction relative to HIII, especially under combined belt and airbag loading, has not been clearly established. The majority of PMHS validation tests have not included airbags. Design requirements focused on pendulum impacts and quasi-static indentor tests [13][15-17]. The 13 sled test restraint conditions used for establishing THOR injury risk curves consisted of 9 standard 3-point belts without airbags, 1 inflatable belt, 1 lap belt and airbag, and two 3-point belts with passenger airbags [18]. Even the last two conditions may not characterize risks to a driver with fundamentally different excursion limits and airbag-packaging restrictions. Also noteworthy is that most of the shoulder belts did not include pretensioners or load limiters and that most lap belts were minimally loaded due to a rigid knee bolster placed in pre-test contact with each subject's knees. A separate assessment of these tests concluded that the proposed criteria "did not predict the risk of rib fractures better than the centre deflection measured on HIII" [19].

In tests that have been conducted with a driver airbag and load limiting seat belt with pretensioner, THOR's response has differed from that of HIII, but not in a way that is clearly more biofidelic. Albert *et al.* [20] compared THOR and HIII with PMHS in two combined restraint and one belt-only condition on a sled. They reported that between the two ATDs, the HIII response was, on average, more similar to the PMHS, better reflecting the shape and peak timing of PMHS deflections measured externally. They also found that the belt-only condition resulted in the highest predicted injury risk for HIII, but the lowest for THOR. They concluded that "more work is needed to evaluate the thoracic biofidelity of the THOR-M under more experimental conditions". Similarly, Forman *et al.* [21] reported different ATD responses to the restraint system in paired full-width 56 km/h crash tests, with THOR riding higher and pushing farther into the airbag than HIII. They observed that restraint system optimization will differ by ATD and vehicle pulse and stated that insufficient PMHS data exist to determine whether either dummy would encourage designs that improve human protection. It should be noted that earlier versions of THOR were stiffer [15] and had responses more similar to HIII [11][22]. On average, HIII deflections in paired tests actually were greater than THOR deflections [23-24], in contrast to more recent comparisons [20-21].

Despite unresolved questions regarding the ability of THOR to assess the thoracic injury risk to occupants restrained by a belt and airbag, the ATD was introduced into European and Australian crash test ratings in 2020 and will be used in China in 2022. NHTSA has indicated repeatedly its intention to do the same in the US [25-26]. The Euro NCAP mobile progressive deformable barrier test includes a THOR 50th percentile male ATD in the driver seat and rates chest protection on the basis of the maximum resultant deflection at any of the four measurement locations (R_{max}). The test protocol notes: "the injury risk data is relevant for seat belt only loading rather than combined seat belt and airbag loading. No change is made in the event of combined seat belt and airbag restraint. This avoids value judgements about the extent of airbag restraint on the chest" [27]. This is not ideal in a test where front airbag deployment is expected.

Improved thoracic injury protection in frontal crashes may be the single most pressing crashworthiness issue

in the passenger vehicle fleet. Perhaps the quickest way to make gains in this area would be the use of a metric in crash test rating programs that is demonstrated to predict field injury risk for drivers restrained by a seat belt and airbag. HIII sternum deflection has demonstrated some limited capability in this regard, but there are concerns associated with placing an even greater emphasis on this single metric. The goal of this study was to determine whether THOR is a better tool than Hybrid III for assessing real-world driver thoracic injury risk under combined restraint loading.

II. METHODS

Vehicle models involved in real-world frontal crashes meeting our inclusion criteria were selected for crash testing with a THOR 50th percentile male ATD in the driver seat position. Using logistic regression, ATD and vehicle metrics collected during these tests and previous THOR and HIII tests were evaluated for their ability to predict thoracic injury outcomes after controlling for specifics of the real-world crash. Two field crash datasets were used in this process. A larger, weighted NASS-CDS sample (n = 902 raw cases) was used to estimate baseline injury odds due to non-vehicle factors. We used a smaller, unweighted sample of NASS-CDS and Crash Investigation Sampling System (CISS) cases (n = 57) with representative THOR and HIII tests to assess the influence of vehicle-related factors on injury outcome. A flow chart of the field injury modeling process is shown in Fig. 2. Below, the study methods specifically related to the injury modeling process are described first (with references to Fig. 2) followed by details of the crash test procedure.

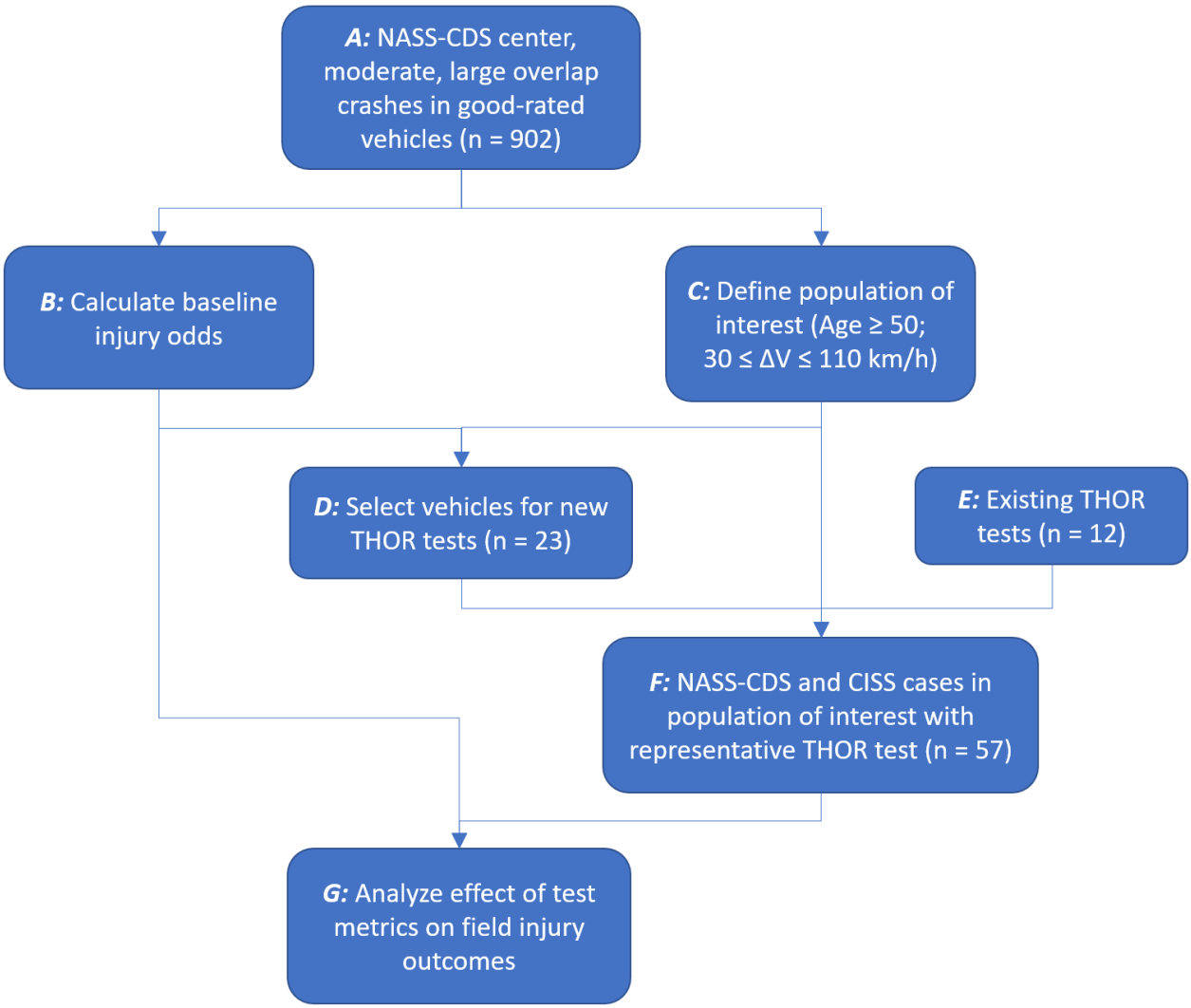

Fig. 2. Modeling field injury outcomes using crash test metrics. Note: ΔV = delta V.

Field crash data: NASS-CDS (Item A in Fig. 2)

Field crash data collected from NASS-CDS and processed as part of two earlier studies by Brumbelow [4][7] served as the foundation of the current study. NASS-CDS was a sample-weighted survey of police-reported crashes in the US conducted by NHTSA from 1979 to 2015. Vehicles with a Good rating in the IIHS moderate overlap test

were included if they were involved in a frontal impact and if the driver was aged 18 or older and restrained by a 3-point seat belt and deployed front airbag. Delta Vs from vehicles equipped with event data recorders (EDRs) were used to calculate "EDR-equivalent" delta Vs for other vehicles using the WinSMASH delta V and the front crash configuration. Crash configurations were assigned based on photographic documentation of damaged vehicle structures. Configurations other than moderate overlap, large overlap, or centre impact were determined to have WinSMASH delta Vs that could not be reliably adjusted and were excluded from further study. There were 902 raw cases that met inclusion criteria.

Calculate baseline injury odds (Item B in Fig. 2)

A wide range of factors contribute to driver injury risk in a specific real-world crash. While some of these are a function of vehicle design and may be quantifiable using an ATD in a test, others are particular to crash and driver characteristics and must be accounted for in order to evaluate the relationship between test metrics and injury when working with a small sample. The effects of delta V, driver age, and driver weight on the risk of AIS≥2 and AIS≥3 thoracic injury were estimated using survey-weighted logistic regression models of the 902 NASS-CDS cases. As discussed by Brumbelow [4], preliminary models also included driver stature and sex, but these were not significant predictors of thoracic injury.

As the effect of age on thoracic injury may not be linear across the range of ages studied [4], preliminary models were fit using natural cubic splines to estimate the effect of age as well as delta V and driver weight. Regression splines allow parameter effects to be estimated using a continuous risk function without assuming a linear relationship with the outcome [34-35]. Because the estimated effects of delta V and driver weight on AIS≥3 injury and driver age on AIS≥2 injury were similar, whether included as linear parameters or cubic splines, the final baseline regression models included these terms as untransformed linear parameters. Splines were calculated using the "splines" package in R [36]. Boundary knots were set at the 5th and 95th quantiles and interior knots at the 35th and 65th quantiles, following the recommendation of Harrell [34], when transition points are unknown beforehand. These knots define the join points of the polynomial functions comprising the overall cubic spline.

Driver weight was missing for 11% of cases used for the baseline models. Multiple imputation enabled inclusion of these cases. The multiple imputation process involves filling in missing data using predictions formed from regressions of the observed data. This imputation is done multiple times for each missing value and the resulting within- and between-imputation variance is used to calculate the uncertainty in the resulting effect estimates. The "mice" package in R [33] was used to impute the missing driver weights 20 times and to pool regression model results. In addition to the parameters included in the injury risk model, driver height and sex were used as predictors for the imputation of driver weight. The R package "survey" [37] was used to fit the baseline models while accounting for case weights in NASS-CDS.

Population of interest and test vehicle selection (Items C and D in Fig. 2)

The full NASS-CDS dataset includes vehicle designs represented by 231 distinct HIII moderate overlap crash tests. As financial and time constraints prohibited conducting the same number of crash tests with THOR, selection of the most relevant case vehicles was first limited to crashes with drivers aged 50 years and older and longitudinal delta Vs of 30–110 km/h. The choice of these criteria was discussed in Brumbelow's 2020 study [7]. Additional factors were considered in prioritizing case vehicles for testing. First, vehicles with IIHS ratings based on verification test results submitted by manufacturers were excluded. Previous analysis of HIII metrics revealed that results from these tests were less predictive of field injury outcomes, and it could not be determined whether this was due to differences between test labs or the vehicles themselves [7]. Second, all case vehicles in which the driver sustained an AIS≥3 thoracic injury were prioritized for testing. Third, cases without an AIS≥3 thoracic injury were prioritized in descending order of the AIS≥3 risk predicted from the baseline injury model described above using the case-specific driver age and delta V, and the mean driver weight. Mean driver weight was used at the vehicle selection stage to allow the inclusion of cases with missing weights and to avoid creating a test dataset with unrepresentative weight values. After considering these factors, 25 vehicles matching one or more of the NASS-CDS case vehicles were selected for testing. One of these vehicles was not readily available for purchase and one of the conducted tests was subsequently excluded due to a late-firing airbag and no pretensioner deployment.

Identify field cases with matching THOR test (Items E and F in Fig. 2)

Some of the 23 vehicle designs tested for this study matched multiple NASS-CDS cases in the population of interest, and these were included in the injury risk analyses. To further increase the field data sample, 2017–2019 CISS files were queried for cases involving the tested vehicle designs. Finally, both NASS-CDS and CISS were queried for cases involving 12 designs previously tested by IIHS with a THOR 50th percentile male driver for a separate study [28]. CISS is NHTSA's replacement of NASS-CDS, and 2017 was the first full year of data collection. For this study, the NASS-CDS inclusion criteria were applied to CISS, including crash configuration assignment based on case photographs. Delta Vs for vehicles without EDRs were adjusted using the configuration-specific regression equations developed earlier [4]. A total of 50 NASS-CDS and 7 CISS cases met inclusion criteria and had matching THOR and HIII crash tests.

Model field injury outcomes (Item G in Fig. 2)

For the 57 field crashes matching tests, logistic regression was used to estimate the effect of different test metrics on the odds of driver AIS≥2 or AIS≥3 thoracic injury after controlling for the risk due to driver age, mass, and delta V. Separate models were specified for each AIS threshold and 14 different vehicle and test metrics. These included metrics from THOR ODB testing, HIII ODB testing, HIII full-width NCAP testing, and the static measurement of each vehicle's bumper-to-firewall distance. HIII metrics consisted of the sternum deflection and thoracic acceleration from both the 64 km/h moderate overlap and 56 km/h full-width test modes. THOR metrics included R_{max}, PC Score, thoracic injury criterion (TIC), and the 3 ms clip of the thoracic acceleration. A correlation analysis of crash test results (Appendix A) was performed to identify other THOR metrics of interest for inclusion in the field injury models. Four additional metrics were selected: maximum resultant deflection at the lower left thorax, peak T12 shear force, maximum T4 rotation, and the sum of the loads from the lap-belt and femur load cells. Finally, the effect of crash test shoulder-belt force on injury risk was estimated. As with prior work [7], shoulder-belt force was characterized using the 20 ms clip (including all peaks) of the tension measured at the upper shoulder-belt. This load was measured in most of the THOR ODB tests and many of the HIII NCAP tests. The regression models utilized the load from either test or the average when both were available. This decision was based on observations of the time histories, which confirmed that both crash test pulses were severe enough to initiate all load limiter stages, and on previous work demonstrating THOR and HIII produce similar measures of upper shoulder belt tension when tested in the same configuration [20-21]. Fig. 2 shows shoulder belt tension values from the tests used in this study. Differences, most of which were small, may be related to the different test configuration and/or seating procedure.

The THOR test vehicle selection process included injury outcome as one of the criteria; many non-injury cases with relatively large weights were not selected for testing. It would be inappropriate to include case weights in the logistic regression models based on test metrics and, consequently, to independently estimate the effects of driver age, driver weight, and delta V. Instead, the baseline injury odds models described above were used to calculate case-specific injury log-odds based on these three parameters. The baseline log-odds were then included as *a priori* offsets in the regression models estimating test metric effects. The Monte Carlo method was used to quantify the uncertainty associated with baseline injury estimates by constructing and sampling from 1,000,000 unique effect estimates for each metric. The Monte Carlo process is detailed in Appendix D. For comparison, results were reported using both the Monte Carlo procedure and the single-value offset applied without accounting for uncertainty in the baseline parameters.

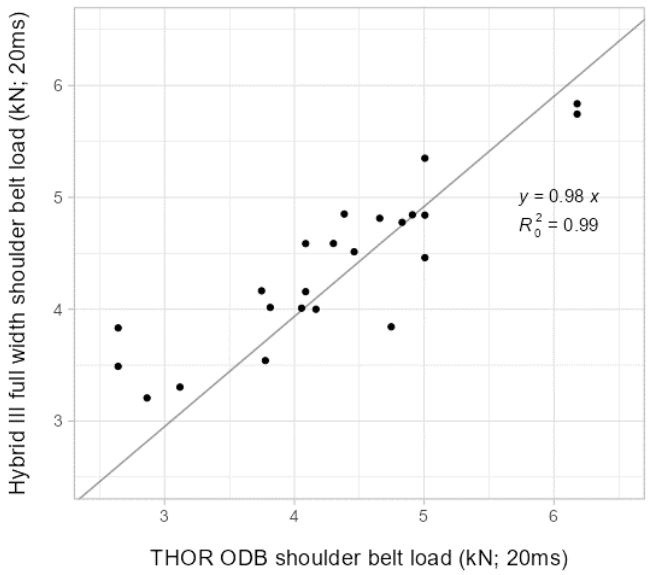

Fig. 2. Shoulder-belt loads in THOR ODB and HIII full-width crash tests matched by vehicle model. Note the regression intercept has been set to 0 and R_0^2 calculated, not R^2.

Three vehicles matched to field crashes were missing shoulder-belt load values, and three were

missing deflection data from one IR-TRACC. In addition, HIII deflection and acceleration data were missing for one vehicle each from the ODB and full-width tests. Among the 57 field cases, 12 were missing driver weight. Multiple imputation was used to account for missing test and field crash data. The "mice" package in R [33] was used to impute the missing data 25 times and to estimate the resulting uncertainty in the model results. In addition to the parameters included in the injury risk models, driver height and sex were used as predictors for the imputation of driver weight. The tension-flexion component of THOR N_{ij} was used as an additional predictor of shoulder-belt force based on results from the correlation analysis (Appendix A).

Thoracic injury outcome was measured using the 2008 version of the AIS scale [38]. NASS-CDS cases that only contained injury codes on the 1995 scale were mapped to the newer scale. Where a single 1995 code could match multiple 2008 codes, case details were sufficient to determine the presence of a 2008 AIS≥2 or AIS≥3 injury.

To compare the predictive ability of different injury metrics, two summary statistics were calculated for each model. The first was based on the area under the receiver operator characteristic curve (AUROC). The AUROC takes a value from 0.5 to 1 and can be interpreted as the probability that the estimated risk for a randomly chosen case with injury will be greater than that for a randomly chosen case without injury. AUROC frequently has been used to assess the ability of regression models to predict binary outcomes, including thoracic injury outcomes from vehicle restraint systems [39]. The AUROC of each model using an ATD or vehicle metric as a predictor was compared with the AUROC for the baseline injury odds model. The average classification improvement (ACI) of each model relative to the baseline model was calculated as:

$$ACI = {(AUROC_{model} - AUROC_{base})}/{(1 - AUROC_{base})} \quad \text{(Eq. 1)}$$

The second summary statistic was based on the Brier score for each model [40]. While a change in AUROC requires one or more pairs of injury and non-injury cases to change order when ranked by predicted injury risk, the Brier score is the mean squared error, so the difference between the predicted risk and the observed outcome (injury or no injury) for each case will contribute to the overall value. It is given as:

$$Brier\ score = \frac{1}{N}\sum_{i=1}^{N}(\hat{p}_i - o_i)^2 \quad \text{(Eq. 2)}$$

where N is the number of NASS-CDS and CISS cases, $\hat{p}_i$ is the predicted injury risk for the ith case and o_i is the observed injury status for the ith case (1 for injury, 0 for no injury). A Brier score of 0 would indicate that a model perfectly predicted all injury outcomes ($\hat{p}$ = 1.0 for all injured drivers and $\hat{p}$ = 0.0 for all non-injured drivers). To compare the performance of each model relative to the baseline, the Brier skill score (BSS) was calculated:

$$BSS = 1 - {Brier\ score_{model}}/{Brier\ score_{base}} \quad \text{(Eq. 3)}$$

For each model calculated using a crash test metric, a negative BSS indicates inferior predictions relative to the baseline, a score of 0 shows no improvement, and scores between 0 and 1 reflect improved predictions.

THOR crash test procedure

All THOR crash tests were 40% overlap, 64 km/h impacts conducted according to the IIHS protocol and seating procedure [29-30], with the following two modifications for THOR positioning. First, the target H-point was adjusted both forward and upward by 20 mm to account for the larger distances between the THOR H-point and pelvis posterior and inferior surfaces. Second, the seat-back angle was adjusted, if necessary, to achieve a pelvic angle of 33°± 2.5° instead of the torso recline angle specified for HIII. The THOR pelvic angle target was consistent with NHTSA's revised THOR-50M seating procedure [31], as was the +9° spine setting.

All tests were conducted using the same THOR ATD, a Standard Build Level A (SBL-A) with the exception of the neck, which had been upgraded to SBL-B. The ATD also was equipped with a modified shoulder pad manufactured by Humanetics to prevent shoulder-belt entrapment between the pad collar and lower neck load cell [32]. A full certification of the ATD was performed prior to testing. Twice during the test series, certification tests were conducted on the head, face, neck, thorax and abdomen. Appendix C lists certification test parameters that did not fall within target corridors. Because of the small magnitude of the failures and the minimal effect they were likely to have on the thoracic response, components failing certification were not replaced. In addition to the certification tests, thoracic pendulum impacts were conducted after every five crash tests to verify consistency in the thoracic deflection response.

III. RESULTS

THOR crash tests

The 35 THOR crash tests are listed in Table E.I in Appendix E, along with any matching real-world tests from NASS-CDS and CISS. Pre-test clearance measurements and summary injury measures also are given in Appendix E. The maximum resultant thoracic deflection (R_{max}) was recorded at the upper right IR-TRACC location in all the tests, with the second greatest resultant deflection usually recorded by the lower right IR-TRACC. R_{max} values ranged from 41 mm to 66 mm, with a median of 53 mm.

Baseline injury models

Baseline AIS≥2 and AIS≥3 thoracic injury odds were modeled on 902 NASS-CDS frontal crashes. The estimated effects are shown in Fig. 4 and Fig. 5. This graphical presentation follows Harrell's recommendation [34] for nonlinear spline models; the coefficient estimates are not interpretable without the underlying spline basis. Effects are shown as differences relative to reference values for each parameter (60 km/h delta V, 60yo for driver age, 75 kg for driver weight). The uniform vertical axes allow comparison of the effect magnitudes across parameters. For linear effects, the slope of the line represents the parameter estimate, while 95% confidence intervals that do not include zero indicate statistical significance at alpha = 0.05. The hash lines in Fig. 4 and Fig. 5 show the delta V, age, and weight values from the smaller sample of 57 drivers with representative THOR tests.

Crash delta V and driver age had stronger effects on the risk of AIS≥3 thoracic injury than on the risk of AIS≥2 thoracic injury. The effect of delta V on AIS≥2 injury exhibited nonlinearity at values below approximately 50 km/h, indicating that the odds of moderate thoracic injury are fairly constant at lower crash severities. In contrast, the odds of serious (AIS≥3) thoracic injury continue to fall with delta V at low severities. The effect of driver age on the risk of AIS≥2 injury did not exhibit nonlinearity, and the linear effect was not statistically significant at alpha = 0.05. For serious injuries, the nonlinear effect of age increased most sharply from around 40yo to 60yo, and then more gradually after that. The effect of driver weight was the weakest contributor to baseline injury risk at both severities. For AIS≥2 injuries, the weight effect was nonlinear, with a minimum estimated injury odds associated with values around the median 75 kg and higher odds at values below and above this weight.

Test metric injury models

Figure 6 shows the delta V, age, and AIS≥3 thoracic injury status for the 57 real-world cases with representative THOR tests, along with the estimated baseline injury risk for delta V, age, and mean driver weight. Of the 57 drivers in the field crashes, 20 sustained an AIS≥3 thoracic injury and 31 sustained an AIS≥2 thoracic injury. The results of regression models including a single crash test or vehicle metric along with baseline log odds offsets for AIS≥2 and AIS≥3 thoracic injury are shown in Tables B.I and B.II, respectively. Results are given using baseline injury odds offsets calculated for each case from the single point (mean) effect estimates for delta V, age and weight and those calculated using the Monte Carlo method to quantify uncertainty associated with the baseline effects. The difference between effect magnitudes calculated using the two methods was usually less than 10%. Confidence intervals generally were wider using the Monte Carlo method, and the change was greater for AIS≥2 than for AIS≥3 models. However, even at the lower injury severity, most resulting p-value differences were minor, demonstrating that uncertainty in the effects of non-vehicle, baseline risk factors was not a major source of uncertainty in effect estimates for vehicle-related factors.

Effects of several crash test or vehicle metrics on AIS≥2 injury outcome were significant at the alpha = 0.05 level using the mean and/or Monte Carlo baseline estimates. Higher injury odds were associated with shorter bumper-to-firewall distance, greater shoulder-belt force, greater THOR resultant deflection at the lower left IR-TRACC, lower THOR T12 shear force, and higher THOR T4 forward rotation. No additional metrics had effects significant at the alpha = 0.1 level. For AIS≥3 injuries, the estimated effects of three metrics were significant at alpha = 0.05 using both the baseline mean and Monte Carlo estimates. Higher injury odds were associated with greater shoulder-belt force, greater THOR resultant deflection at the lower left IR-TRACC, and lower THOR R_{max} (which occurred at the upper right IR-TRACC in all tests). Additional metrics with associated p-values ≤ 0.1 were THOR T4 rotation and TIC.

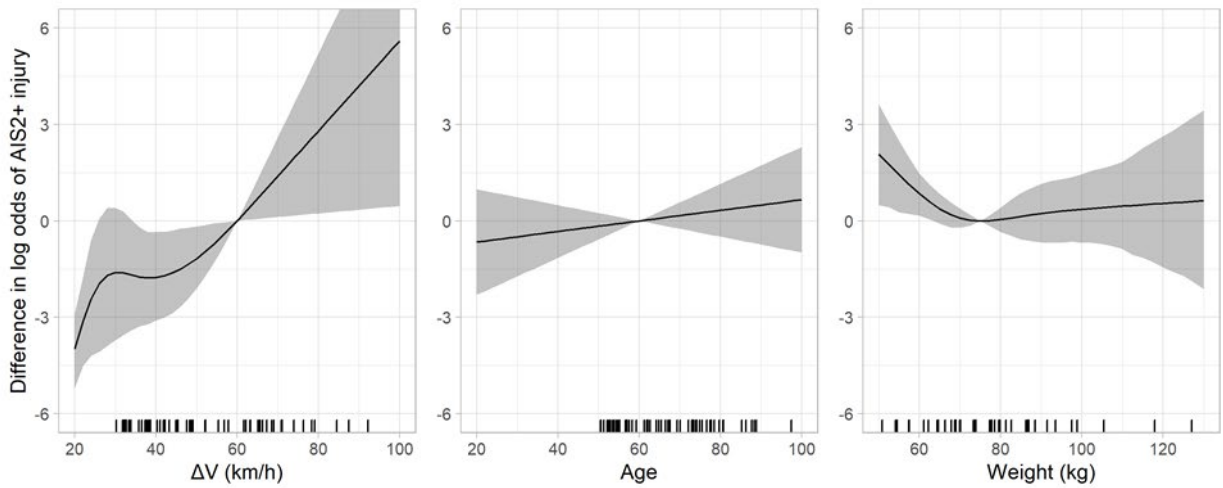

Fig. 4. Baseline effects of delta V (ΔV), age, and weight on log odds of AIS≥2 thoracic injury relative to a 60yo, 75 kg driver in a crash with a 60 km/h ΔV. Effects based on results of weighted model of 902 NASS-CDS cases. Shaded area represents the 95% confidence interval. Hash lines indicate values for 57 cases with a representative THOR test.

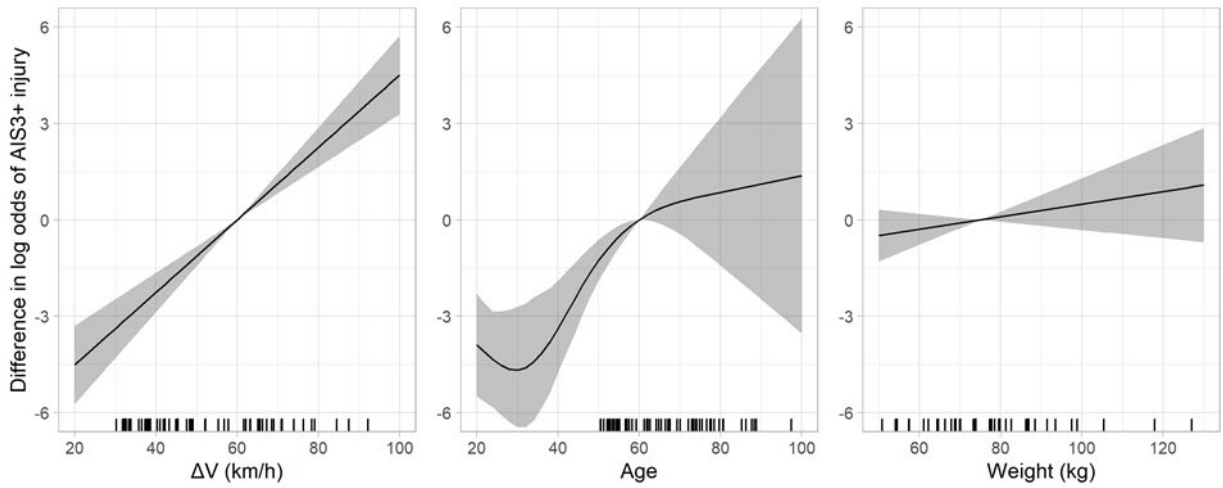

Fig. 5. Baseline effects of delta V (ΔV), age, and weight on log odds of AIS≥3 thoracic injury relative to a 60yo, 75 kg driver in a crash with 60 km/h ΔV. Effects based on results of weighted model of 902 NASS-CDS cases. Shaded area represents the 95% confidence interval. Hash lines indicate values for 57 cases with a representative THOR test.

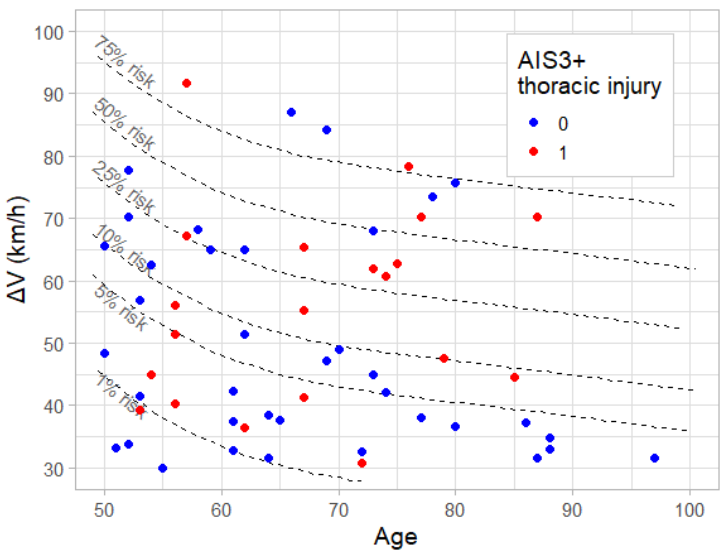

Fig. 6. Crash delta V (ΔV), driver age, and AIS≥3 thoracic injury outcome for the 57 NASS-CDS and CISS cases. Dashed lines indicate baseline injury risks at the median driver weight.

The prediction performance of models including one vehicle or test metric was compared with the baseline models using only delta V, age and weight. The average classification improvement (ACI, Eq. 1) is a measure of the relative injury risk ranking of injury and non-injury cases, while the Brier skill score (BSS, Eq. 3) measures the difference between predicted risk and injury outcome for each case. For AIS≥2 injury, ACI and BSS values produced similar rankings of model performance. Injury models based on THOR T12 shear force and THOR T4 rotation improved predictions most, with shoulder-belt force, bumper-to-firewall distance and THOR lower left thoracic deflection also showing greater than average improvements. For AIS≥3 injury, there was somewhat less

agreement between ACI and BSS, and model performance was less evenly distributed. Shoulder-belt force was the best performing metric and THOR lower left thoracic deflection the second best, but there was a relatively large difference between these models. The remaining metrics demonstrated little to no improvement in injury prediction as measured by ACI and/or BSS.

Figure 7 compares the THOR R_{max} and upper deflection differences in the ODB tests with previously published sled test results. Many of the ODB deflection differences were lower than belt-only sled tests with similar R_{max} values, but there was still a large degree of overlap. It is also noteworthy that the R_{max} values in the ODB tests were similar to results from sled tests with much lower delta Vs. Most of the sled tests used in development of THOR's injury risk functions had shorter crash pulses, no belt pretensioning or force limiting, and a rigid knee bolster placed at initial contact with the ATD that resulted in minimal lap-belt load [14]. Each of these factors has the potential to affect the relationship between THOR rib deflections and injury risk. The ODB tests conducted for this study produced a median R_{max} value of 53 mm, which corresponds to an 86% risk of AIS≥3 thoracic injury for a 60yo using the sled test-derived risk curve [41]. In contrast, the baseline field data model estimates a 38% risk in a 70 km/h delta V crash (95% CI: 10-74%).

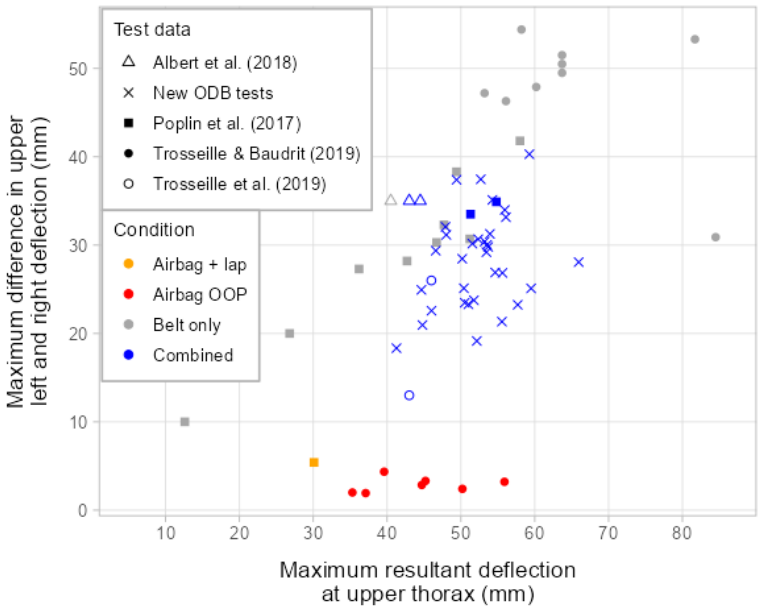

Fig. 7. THOR resultant and differential deflections for ODB crash tests compared with previously published data by restraint condition. One inflatable belt test is not shown. OOP = out of position.

IV. Discussion

For this set of matched field crashes and crash tests, proposed THOR injury metrics were unable to predict real-world thoracic injury outcomes better than measurements taken from HIII or directly from the vehicle, such as shoulder-belt force and bumper-to-firewall distance. In fact, there is evidence that R_{max}, the primary THOR metric in use today [27][41], is inversely related to serious injury outcomes, with higher crash test values associated with significantly lower likelihood of AIS≥3 injury and lower likelihood of AIS≥2 injury. This result, when taken together with the results from other models of field injury, suggests THOR may not represent human thoracic response to combined belt and airbag loading.

Models including shoulder-belt force were among the best performing predictors of both AIS≥2 and AIS≥3 injury. If THOR R_{max} was simply a function of shoulder-belt force, it would also provide directionally correct injury risk predictions. Linear regression models of crash test data (Table A.I) indicated that shoulder-belt force was a significant contributor to R_{max}, but the correlation between these two metrics was relatively weak (R = 0.48, Fig. A.1) and the regression models suggested that loading from the rest of the restraint system influenced deflection values as well. In general, this is to be expected and reflects human loading; holding the shoulder-belt load constant while increasing loads from other sources will increase deflection. However, the ATD must respond similarly to a human driver to the balance between restraint loads, and not only when considering peak values but as a function of time. It was not possible to directly measure non-belt restraint forces in the THOR crash tests, but R_{max} models (Table A.I) showed that, after controlling for belt load, THOR deflections tended to be lower for vehicles that allowed more torso rotation, more negative T12 shear forces, or lower combined femur and lap-belt loads, all of which could indirectly reflect lower overall restraint from the airbag. ("Airbag" here is used to refer to the combined loading from the airbag and steering column.) Yet the results of field injury models using torso rotation and T12 shear force were opposite, suggesting that vehicles with lower THOR rotation and more positive shear forces had lower risk of AIS≥2 (p-values of 0.02–0.04) and AIS≥3 (p-values of 0.11–0.12) injury. Along with shoulder-belt force, these metrics provided the greatest improvement in predicting AIS≥2 outcomes. This suggests that a greater contribution of loading from the airbag is beneficial for human drivers, even as it increases

THOR deflection.

An additional, though related, observation further indicates that THOR rib compression may be overly sensitive to the relative contribution of airbag loading in the restraint system. Increasing resultant thoracic deflection at the lower left IR-TRACC did predict increased injury risk, but the three-dimensional rib deflection at this location is unique. The lower left ATD thorax was not loaded directly by the belt and the rib bulged outward, producing positive deflection in the x-direction. This behaviour has been observed previously [14][23][42]. The degree of bulging and consequent resultant deflection appeared to be partially governed by the degree of airbag loading: vehicles that produced more torso rotation and more negative T12 shear force, presumably from lower airbag loads, tended to have greater resultant deflection at the lower left IR-TRACC (Fig. A.1, also confirmed with models controlling for belt force). Again, field crash models indicated that moving the restraint balance in this direction was more likely to cause injury; a deflection increase of around 5 mm at the lower left IR-TRACC was associated with an AIS≥2 odds ratio of 2.7 ($p = 0.04$) and an AIS≥3 odds ratio of 3.5 ($p = 0.01$). While anterior rib displacement has been observed and discussed as a possible injury mechanism in PMHS tests [43], THOR's biofidelity in this regard is an open question given the focus on posterior deflection during dummy development. Shaw *et al.* [13] demonstrated that THOR exhibited almost no coupling between this location and the rest of the rib cage, compared with around 50% coupling for PMHS, but a later study [14] suggested that PMHS bulging of the lower thorax is much greater than THOR's due to inertial loading of the internal organs. It is possible that THOR anterior rib deflection is a general marker of restraint system characteristics that influence injury in other ways.

The limitations of R_{max} as a tool to assess the risk associated with a variety of restraint systems has motivated work to establish metrics incorporating the differential deflection from multiple locations on the thorax. PC Score and TIC are two examples. In this study, neither metric improved the prediction of AIS≥2 or AIS≥3 injury. As both metrics include R_{max}, this is not surprising. However, separate analyses indicated that the difference between the upper THOR deflections did not improve injury prediction on its own. Furthermore, the magnitude of the deflection difference had a similar relationship as R_{max} to torso rotation and T12 shear force. In other words, for this set of crashes, the difference in upper left and right deflection did not appear to improve the dummy's ability to distinguish between the relative contributions of the airbag and belt.

Comparing the individual IR-TRACC readings from this test series with sled test data reported by Albert *et al.* [20] suggests a reason that THOR deflection differences are opposite of what would be expected for belt and airbag loading. Unlike the ODB test series, the sled tests isolated the effect of restraint changes on the dummy's thoracic deflection. The addition of an airbag to a belt-only restraint resulted in an increase of around 4 mm of resultant deflection at the upper right IR-TRACC, but a slight decrease (2–3 mm) at the upper left. A third restraint condition that included a knee-bolster airbag along with the steering wheel airbag produced similar deflection at the upper right, but another 5–10 mm decrease at the upper left. The component-level deflections reveal that the upper left posterior deflections were similar for all conditions, but that the resultant was driven primarily by upward vertical deflection of the rib, which decreased in the airbag conditions. Similarly, in the ODB crash tests, upward movement measured by the upper left IR-TRACC was often the largest component at this location and the amount of vertical movement was correlated with the shoulder-belt load. In both the sled tests and the ODB crash tests, the vertical movement at the upper left location peaked after the maximum longitudinal deflection at the upper left and right. The maximum differences between upper right and left resultants were highly sensitive to the offset between these peaks.

While the TIC calculation includes only upper deflection differences, PC Score also includes differences measured at the lower thorax. However, the order of operations means that only resultant values are compared; anterior and posterior deflections of the same magnitude cancel each other. In the ODB tests, the increase in the three-dimensional separation between the upper thorax measurement points was 18% greater than the maximum difference in resultant deflections, on average. The corresponding value at the lower thorax was 47%, largely due to the anterior bulging at the lower left sensor location.

If THOR is unable to accurately measure injury risk under combined restraint loading, one possible explanation is the shoulder/clavicle design of the ATD. Others have demonstrated that the forward movement of the shoulder during excursion is not biofidelic [14][44]. As the shoulder/clavicle assembly moves forward relative to the spine, it reduces shoulder-belt loading of the ribs. Because of this, and in addition to the effect of tighter belt wrap around the shoulder, greater amounts of ATD rotation and excursion also could allow time for more shoulder excursion, further reducing belt loading of the thorax. If the shoulder behaviour has less effect on airbag loading

of the ribs, any forward movement shifts the rib deflection source more towards the airbag. Relative to a human, this would be expected to produce both a reduction in the ATD's sensitivity to belt loading and an increase in its sensitivity to airbag loading. NHTSA research on an alternate THOR shoulder design is ongoing [26].

The results of this study are consistent with previous research in reinforcing the need for THOR validation against PMHS responses to combined belt and airbag loading. Hu *et al.* [45] evaluated restraint technologies in rear seat sled tests and found that reducing belt loads and adding an airbag or additional shoulder belt caused THOR deflections to increase, in contrast with the deflections of HIII 5th percentile female and 95th percentile male ATDs. Petitjean *et al.* [42] found that a combined restraint system with a 4 kN shoulder-belt restraint produced similar THOR R_{max} values as a 6 kN belt-only restraint, despite being associated with reduced injury outcomes for PMHS and living humans [46]. Trosseille *et al.* [47] constructed two combined restraint systems, one more belt-intensive and the other more airbag-intensive, and then applied a unique sled pulse for each system to produce a similar THOR R_{max}. They reported large differences in sensitivity to the restraint and pulse changes between PMHS and THOR, with PMHS exhibiting higher head rotation, velocity and excursion with the belt-intensive condition, while THOR showed similar head rotation and reductions in velocity and excursion. While shoulder-belt loads were similar for the two surrogates due to the force limiter, airbag forces were 40–60% greater for THOR in both conditions and lap-belt loads were 50–100% greater, with substantial differences remaining after scaling for mass. Davidsson *et al.* [48] previously had reproduced PMHS tests with THOR and showed the risk associated with R_{max} was restraint-dependent. This observation was long associated with HIII sternum deflection [8][5][11] and represented potentially the greatest opportunity for THOR to improve serious injury risk prediction. This study serves as additional evidence that THOR has not filled this gap.

None of the HIII metrics was a consistent predictor of field injury outcomes. Sternum deflection measured in the ODB configuration was the best performing HIII metric when considering AIS≥2 injuries, but this did not hold for AIS≥3 injuries. As sternum fractures did compose the majority of the AIS2 injuries in this dataset, HIII deflection in the ODB test may be a valid indicator of injury at the measurement location. Fully investigating this possibility would require a separate model of sternum fractures, including any experienced by drivers who also sustained an AIS≥3 injury. This was beyond the scope of this study.

In lieu of an ATD deflection-based measure able to predict injuries of all severities, crash test ratings that consider shoulder-belt force may have greater relevance to field crashes. The Euro NCAP driver chest rating is based on THOR R_{max} but also includes a penalty for belt loads exceeding 6 kN [27]. It should be noted that the wide range of vehicle model years in this study produced a mean belt load (4.5 kN sustained for 20 ms) that may be higher than typical for the modern fleet. The implementation of shoulder-belt loads into rating criteria should account for the likelihood that the relationship between belt load limit and injury risk is nonlinear. At the lowest belt forces, injury risks may increase due to greater excursion, secondary impacts, or even insufficient pretensioning. And despite the high performance of shoulder-belt force in predicting injury outcomes relative to most other metrics, the results for torso rotation and T12 shear force (which were not correlated with shoulder-belt force in the test data) imply that a more comprehensive injury metric is possible given the right assessment tools.

Crash test metrics aside, the baseline injury risk models are themselves helpful for characterizing thoracic injury risk in field crashes. There were notable shifts in the relative injury log-odds curves for delta V, age and weight depending on whether AIS2 injuries were assessed (Fig. 4 and Fig. 5). In general, baseline effects were weaker predictors of AIS≥2 injuries, suggesting that individual driver injury tolerances or crash factors not captured by delta V may be more influential in moderate than in serious injury outcomes.

Limitations

The nonlinear relationship between crash delta V and AIS≥2 injury outcome illustrates one limitation of this study. All THOR crash tests were conducted at the same 64 km/h impact speed, which produces a delta V of approximately 70 km/h. These tests were used to represent field crashes with a range of ΔVs, under the assumption that test measurements represent a scalable characterization of the vehicle response. However, it is likely that both the relationship between crash test speed and R_{max}, for example, as well as the relationship between R_{max} and injury risk, exhibit nonlinearity to some degree. Accounting for these possibilities would require speed-matched crash testing, which is beyond the scope of this research and consumer-information rating programs. It also seems unlikely that the relationship between R_{max} and delta V would rank the vehicles in this study differently enough to reverse the observed inverse relationship with serious injury outcome.

Still, it is possible that THOR's response would better predict injury outcomes at lower crash severities. The median delta V for drivers with an AIS≥3 injury in the field crashes was around 55 km/h, both in the baseline weighted dataset (n = 902) and in the unweighted dataset of THOR-tested vehicles (n = 57). The sensitivity of the ATD to different types of restraint loading may differ at lower speeds. While limited to belt-only loading, Parent *et al.* [49] reported that the biofidelity rank order of THOR and HIII depended on whether 30 km/h or 40 km/h tests were used. Similarly, the observations made in the 40% overlap scenario may not generalize to other test modes, especially those with more oblique loading.

The small sample size is another limitation of the study. With only 57 field cases, the differences in driver, vehicle and crash factors could still be substantial even after controlling for delta V, age and weight. PMHS datasets used for THOR validation contain similar numbers of observations (e.g. n = 44 in Poplin *et al*.'s study [18]) but maintain stricter, and measurable, controls on most factors relevant to injury outcome. The main advantages of the field data are that they involve living humans and that none of the controls are artificial. The seemingly ideal dataset for ATD validation would be field crash data with counts sufficient to rule out covariance between uncontrolled injury factors and representative dummy responses. In the absence of this ideal, more realistic PMHS tests and/or computational modeling may be required to validate THOR's response to combined belt and airbag loading.

The application of baseline injury odds as *a priori* offsets to the models estimating test metric effects represents a novel approach to studying field crash data. This was necessary to enable the selection and analysis of a subsample of NASS-CDS and CISS cases, but it carries its own limitations. Primarily, it assumes not only that there is no covariance between test metrics and uncontrolled injury factors, as mentioned above, but also that there is no covariance between test metrics and the factors that are controlled. For example, if older drivers tend to drive vehicles with greater bumper-to-firewall distances, and if more crush space reduces injury risk, then the true effect of age would be even greater than that estimated by the baseline models. This, in turn, would result in biased estimates for the test metrics. It was possible to assess covariance between baseline metrics and bumper-to-firewall distance, as well as the IIHS ODB and NCAP metrics, since representative measures existed for the full dataset of 902 cases. All metrics exhibited minimal covariance, with those in the example above, driver age and crush space, demonstrating the strongest correlation with an R value of 0.1. Covariance between THOR test and baseline case metrics cannot be assessed, given the lack of test data for the larger dataset. While it cannot be ruled out, there is no reason to believe it would be greater than measurements taken in crash tests using HIII.

V. CONCLUSIONS

Shoulder-belt force improved predictions of both AIS≥2 and AIS≥3 driver injury in this set of field and ODB laboratory crashes matched by vehicle design. HIII sternum deflection improved prediction of AIS≥2 but not AIS≥3 injuries. Among thoracic injury metrics proposed for THOR, R_{max} was inversely related to AIS≥3 injury outcome. THOR metrics designed to distinguish between degrees of shoulder-belt and airbag load sharing did not appear able to do so. Modeling injury using crash test metrics predictive of R_{max} while controlling for shoulder-belt force suggested that greater airbag loading reduced injury risk even as it increased R_{max}. These results are consistent with published literature in highlighting the need for THOR validation in realistic restraint conditions involving both a shoulder belt and airbag. It is possible that using the dummy in its current form, at least in the 64 km/h ODB test, will encourage vehicle restraint systems that increase thoracic injury risk in certain types of frontal crashes.

VI. ACKNOWLEDGEMENTS

Honda R&D Americas generously loaned the THOR used in all the ODB crash tests, while Humanetics provided valuable technical support. The authors are grateful to both organizations and to the crash test team at the Vehicle Research Center for making this study possible.

VII. References

[1] National Highway Traffic Safety Administration (2020) Fatality Analysis Reporting System.

[2] Forman, J., Poplin, G. S., *et al*. (2019) Automobile injury trends in the contemporary fleet: Belted occupants in frontal collisions. *Traffic Injury Prevention*, **20**(6): pp.607–612.

[3] Brumbelow, M. L., Jermakian, J. S. (2022) Injury risks and crashworthiness Benefits for females and males: Which differences are physiological? *Traffic Injury Prevention*, **23**(1): pp.11–16.

[4] Brumbelow, M. L. (2019) Front crash injury risks for restrained drivers in good-rated vehicles by age, impact configuration, and EDR-based delta V. *Proceedings of IRCOBI Conference*, 2019, Florence, Italy.

[5] Mertz, H. J., Prasad, P., Irwin, A. L. (1997) Injury risk curves for children and adults in frontal and rear collisions. SAE Technical Paper, 973318.

[6] Laituri, T. R., Prasad, P., Sullivan, K., Frankstein, M., Thomas, R. S. (2005) Derivation and evaluation of a provisional, age-dependent, AIS3+ thoracic injury risk curve for belted adults in frontal impacts. SAE Technical Paper 2005-01-0297.

[7] Brumbelow, M. L. (2020) Can front crash rating programs using Hybrid III predict real-world thoracic injuries? *Proceedings of IRCOBI Conference*, 2020, Munich, Germany (postponed).

[8] Horsch, J. D., Melvin, J. W., Viano, D. C., Mertz, H. J. (1991) Thoracic injury assessment of belt restraint systems based on Hybrid III chest compression. SAE Technical Paper, 912895.

[9] Digges, K., Dalmotas, D., Prasad, P., Mueller, B. (2017) The need to control belt routing for silver NCAP ratings. *Proceedings of the International Technical Conference on the Enhanced Safety of Vehicles*, 2017, Detroit, USA.

[10] Brumbelow, M. L. (2020) Adjusting for the effect of shoulder belt placement on Hybrid III sternum deflection. *Proceedings of IRCOBI Conference*, 2020, Munich, Germany (postponed).

[11] Kent, R., Lessley, D., Shaw, G., Crandall, J. (2003) The utility of Hybrid III and THOR chest deflection for discriminating between standard and force-limiting belt systems. *Stapp Car Crash Journal*, **47**: pp.267–297.

[12] Eggers, A., Eickhoff, B., Dobberstein, J., Zellmer, H., Adolph, T. (2014) Effects of variations in belt geometry, double pretensioning and adaptive load limiting on advanced chest measurements of THOR and Hybrid III. *Proceedings of IRCOBI Conference*, 2014, Berlin, Germany.

[13] Shaw, G., Lessley, D., Kent, R., Crandall, J. (2005) Dummy torso response to anterior quasi-static loading. *Proceedings of the International Technical Conference on the Enhanced Safety of Vehicles*, 2005, Washington, DC, USA.

[14] Shaw, G., Lessley, D., Ash, J., Crandall, J., Parent, D. (2013) Response comparison for the Hybrid III, THOR mod kit with SD-3 shoulder, and PMHS in a simulated frontal crash. *Proceedings of the International Technical Conference on the Enhanced Safety of Vehicles*, 2013, Seoul, South Korea.

[15] Parent, D. P., Craig, M., Ridella, S. A., McFadden, J. D. (2013) Thoracic biofidelity assessment of the THOR mod kit ATD. *Proceedings of the International Technical Conference on the Enhanced Safety of Vehicles*, 2013, Seoul, South Korea.

[16] Rangarajan, N., White, Jr. R., *et al*. (1998) Design and performance of the THOR advanced frontal crash test dummy thorax and abdomen assemblies. *Proceedings of the International Technical Conference on the Enhanced Safety of Vehicles*, 1998, Windsor, Canada.

[17] General Engineering and Systems Analysis Company (2005) Biomechanical response requirements of the THOR NHTSA advanced frontal dummy. Report No: GESAC-05-03.

[18] Poplin, G. S., McMurry, T. L., *et al*. (2017) Development of thoracic injury risk functions for the THOR ATD. *Accident Analysis & Prevention*, **106**: pp.122–130.

[19] Trosseille, X., Baudrit, P. (2019) Updated chest injury criterion for the THOR dummy. *Proceedings of the International Technical Conference on the Enhanced Safety of Vehicles*, 2019, Eindhoven, Netherlands.

[20] Albert, D. L., Beeman, S. M., Kemper, A. R. (2018) Assessment of thoracic response and injury risk using the Hybrid III, THOR-M, and post-mortem human surrogates under various restraint conditions in full-scale frontal sled tests. *Stapp Car Crash Journal*, **62**: pp.1–65.

[21] Forman, J., Acosta, S. M., Bollapragada, V., Gepner, B., Wenicke, P. (2019) Comparing restraint system sensitivity between the THOR and the Hybrid III, and potential implications in restraint optimization. *Proceedings of the International Technical Conference on the Enhanced Safety of Vehicles*, 2019, Eindhoven, Netherlands.

[22] Xu, L., Jensen, J., *et al*. (2000) Comparative performance evaluation of THOR and Hybrid III. SAE Technical Paper, 2000-01-0161.

[23] Shaw, G., Crandall, J., Butcher, J. (2000) Biofidelity evaluation of the THOR advanced frontal crash test dummy. *Proceedings of IRCOBI Conference*, 2000, Montpellier, France.
[24] Mueller, B. C., Sherwood, C. P., Arbelaez, R. A., Zuby, D. S., Nolan, J. M. (2011) Comparison of Hybrid III and THOR dummies in paired small overlap tests. *Stapp Car Crash Journal*, **55**: pp.379–409.
[25] Saunders, J., Craig, M., Parent, D. (2012) Moving deformable barrier test procedure for evaluating small overlap/oblique crashes. SAE Technical Paper, 2012-01-0577.
[26] National Highway Traffic Safety Administration (2021) NHTSA safety research portfolio public meeting: Fall 2021; Advanced crash test dummies. 21 October 2021. Webinar.
[27] Euro NCAP (2020) Assessment protocol: Adult occupant protection. Version 9.1.2. 2020.
[28] Jermakian, J., Edwards, M., Jagtap, S. (2020) Development of a crashworthiness evaluation program for rear-seated occupants in frontal crashes. SAE Government / Industry meeting, 2020, Washington, DC, USA.
[29] Insurance Institute for Highway Safety (2017) Moderate overlap frontal crashworthiness evaluation crash test protocol (Version XVIII).
[30] Insurance Institute for Highway Safety (2004) Guidelines for using the UMTRI ATD positioning procedure for ATD and seat positioning (Version V).
[31] National Highway Traffic Safety Administration (2019) Revised THOR 50th percentile male dummy seating procedure. DOT HS 812 746.
[32] Wang, Z. J., Fu, S., McInnis, J., Arthur, J. (2019) Evaluation of novel designs to address the shoulder-belt entrapment for THOR-50M ATD. *Proceedings of IRCOBI Conference*, 2019, Florence, Italy.
[33] van Buuren, S., Groothuis-Oudshoorn, K. (2011) Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, **45**(3): pp.1–67.
[34] Harrell Jr, F. E. (2015) Regression Modeling Strategies, 2nd edition. Springer.
[35] Forman, J., McMurry, T. L. (2018) Nonlinear models of injury risk and implications in intervention targeting for thoracic injury mitigation. *Traffic Injury Prevention,* **19**(S2): pp.S103–S108.
[36] R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, 2021, Vienna, Austria.
[37] Lumley, T. (2004) Analysis of complex survey samples. *Journal of Statistical Software*, **9**(1): pp.1–19.
[38] Association for the Advancement of Automotive Medicine (2008) The abbreviated injury scale-2005 revision, update 2008. AAAM, Des Plaines, USA.
[39] Kent, R., Patrie, J., Benson, N. (2003) The Hybrid III dummy as a discriminator of injurious and non-injurious restraint loading. *Annual Proceedings of the Association for the Advancement of Automotive Medicine*, 2003.
[40] Goldstein-Greenwood, J. (2001) A brief on brier scores. URL: data.library.virginia.edu/a-brief-on-brier-scores.
[41] Craig, M., Parent, D., *et al*. (2020) Injury criteria for the THOR 50th male ATD. National Highway Traffic Safety Administration, 2020.
[42] Petitjean, A., Lebarbe, M., Potier, P., Trosseille, X., Lassau, J. P. (2002) Laboratory reconstructions of real-world frontal crash configurations using the Hybrid III and THOR dummies and PMHS. *Stapp Car Crash Journal*, **46**: pp.27–54.
[43] Shaw, G., Parent, D., *et al.* (2009) Impact response of restrained PMHS in frontal sled tests: Skeletal deformation patterns under seat belt loading. *Stapp Car Crash Journal*, **53**: pp.1–48.
[44] Lemmen, P., Been, B., *et al*. (2013) An advanced thorax-shoulder design for the THOR dummy. *Proceedings of the International Technical Conference on the Enhanced Safety of Vehicles*, 2013, Seoul, South Korea.
[45] Hu, J., Reed, M.P., Rupp, J.D. (2017) Optimizing seat belt and airbag designs for rear seat occupant protection in frontal crashes. *Stapp Car Crash Journal*, **61**: pp.67–100.
[46] Foret-Bruno, J. Y., Trosseille, X., *et al*. (2001) Comparison of thoracic injury risk in frontal car crashes for occupant restrained without belt load limiters and those restrained with 6 kN and 4 kN belt load limiters. *Stapp Car Crash Journal*, **44**: pp.205–224.
[47] Trosseille, X., Petit, P., Uriot, J., Potier, P., Baudrit, P. (2019) Assessment of several THOR thoracic injury criteria based on a new postmortem human subject test series and recommendations. *Stapp Car Crash Journal*, **63**: pp.291–305.
[48] Davidsson, J., Carroll, J., *et al.* (2014) Development of injury risk functions for use with the THORAX demonstrator; an updated THOR. *Proceedings of IRCOBI Conference*, 2014, Berlin, Germany.
[49] Parent, D. P., Craig, M., Moorhouse, K. (2017) Biofidelity evaluation of the THOR and Hybrid III 50th percentile male frontal impact anthropomorphic test devices. *Stapp Car Crash Journal*, **61**: pp.227–276.

VIII. APPENDICES

APPENDIX A:

THOR METRIC CORRELATION ANALYSIS

The relationship between HIII and THOR injury metrics for the thorax and other body regions in the 35 matched pairs of ODB crash tests will be the subject of future work. For the current study, the correlation between various THOR metrics recorded in the tests was analysed. The purpose of the analysis was to evaluate the relationship between different metrics and to identify which should be included in the field injury models.

Pearson correlation coefficients were calculated for three deflection-based metrics that have been offered as indicators of thoracic injury risk along with other measures that may indicate differences in restraint system loads between vehicles. The three previously proposed injury metrics were maximum resultant deflection (R_{max}), PC Score, and thoracic injury criterion (TIC) [18-19]. The other included metrics were shoulder-belt force, lap-belt force, thoracic acceleration and rotation (both measured at T4), thoracic spine shear force (T12), HIC_{15}, the tension-flexion component of N_{ij}, the left and right femur loads, and the sum of the lap-belt and femur loads.

Shoulder-belt force was previously found to be a predictor of injury [7]. Any potential difference in injury prediction performance between shoulder-belt force and R_{max} could be a result of non-belt restraint loads on the dummy. Linear regression models were constructed to further explore the relationship between R_{max} and non-deflection metrics while controlling for shoulder-belt force. Interactions between shoulder-belt force and the other metrics on the effect of R_{max} also were evaluated. Six of the 35 tests were missing shoulder-belt force, three were missing lap-belt force, and three were missing IR-TRACC data from one of the deflection locations. Multiple imputation was used to permit inclusion of these tests in the regression models. Missing data were imputed 20 times and model results were pooled using the "mice" package in the R programming language [33].

Pearson correlation coefficients between THOR metrics are shown in Fig. A.1. The two resultant deflections from the right side of the thorax were strongly correlated with each other and with PC Score. Shoulder-belt force

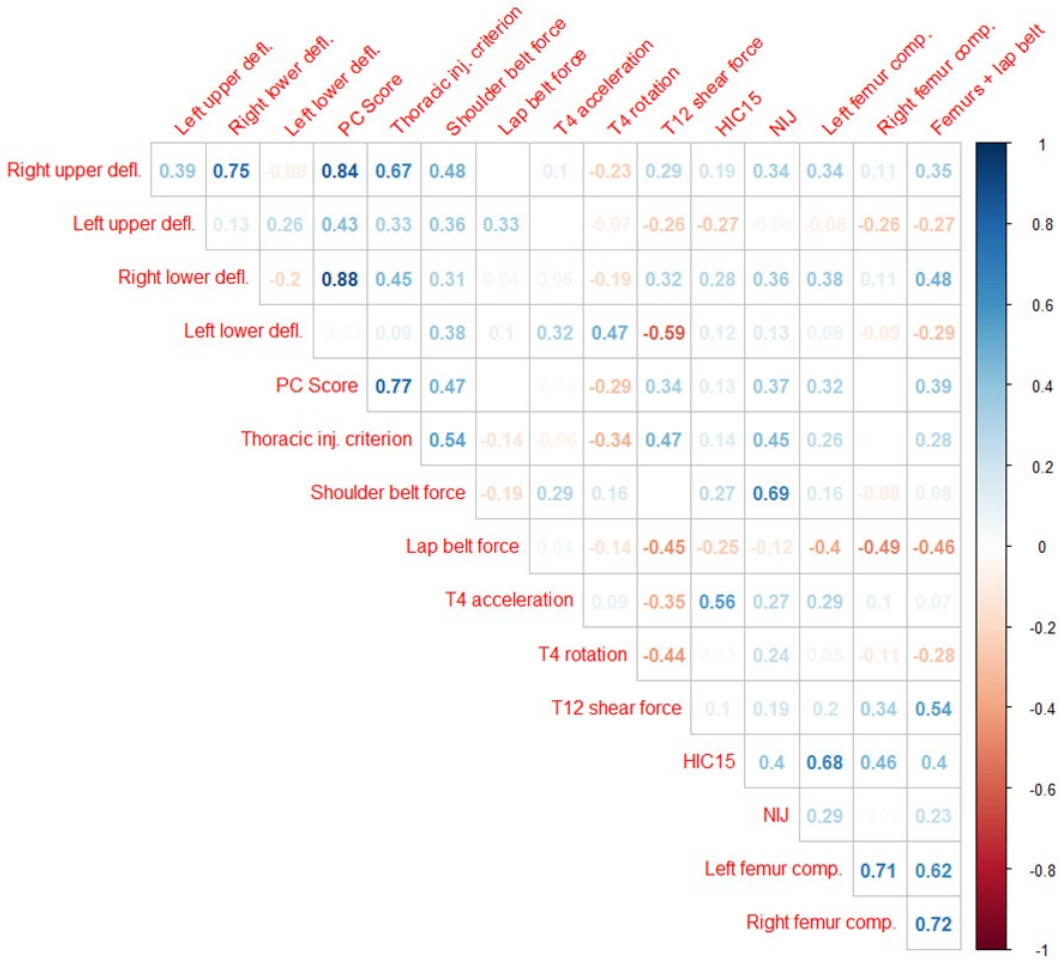

Fig. A.1. Pearson correlation coefficients for THOR metrics measured in 35 crash tests.
comp. = compression; defl. = deflection; inj = injury.

was positively but weakly correlated with all thoracic deflection-based metrics and more strongly correlated with N_{ij} (specifically the tension-flexion component).

Results of linear models characterizing the thoracic response in terms of R_{max} are shown in Table A.I. The main effect of shoulder-belt force was significant at the alpha = 0.05 level in each model and indicated that a 0.9 kN increase in force was associated with an increase in R_{max} of around 3 mm. The shoulder-belt interactions with T4 rotation and combined lap-belt and femur load were also significant at the alpha = 0.05 level. Parameter estimates with p-values from 0.05 to 0.10 included the main effects of T4 rotation, T12 shear force, and combined lap-belt and femur loads, as well as the shoulder-belt interaction with T12 shear force. The main effects of T4 rotation and T12 shear force indicated that after controlling for shoulder-belt load, less additional torso restraint produced lower R_{max} values. The interaction effects for these metrics showed increased shoulder-belt loads had a greater effect on R_{max} when there was less additional torso restraint. The combined lap-belt and femur load models indicated that additional lower-body restraint corresponded to increased R_{max} at the same shoulder-belt load (main effects model), or to a reduction in the effect of increasing shoulder-belt loads on R_{max} (interaction model).

In addition to Rmax, PC Score, and TIC, four additional THOR metrics were chosen for inclusion in the injury prediction models based on the results of this analysis. T12 shear force, T4 rotation, and the combined femur and lap belt load were selected based on their correlation with R_{max} differences that exist when stratifying by shoulder-belt force (Table A.I). Lower left thorax deflection was selected because it was not correlated with Rmax, PC Score, or TIC (Fig. A.1). It also had the strongest correlation with T4 rotation and T12 shear force among all four of the IR-TRACC locations.

TABLE A.I
R_{MAX} LINEAR MODEL RESULTS

		Main effects model		**Interaction model**	
Term	IQR	Estimate	p-value	Interaction effect estimate	p-value
Shoulder-belt force	0.93 kN (20 ms)	3.1	**0.004**	−1.6	0.17
Lap-belt force	1.2 kN (20 ms)	0.6	0.42		
Shoulder-belt force	0.93 kN (20 ms)	2.9	**0.01**	1.7	0.19
T4 acceleration	7.5 g (3 ms res.)	−0.01	1.00		
Shoulder-belt force	0.93 kN (20 ms)	3.2	**0.002**	2.8	**0.03**
T4 rotation	15 deg (forward)	−2.4	0.06		
Shoulder-belt force	0.93 kN (20 ms)	2.9	**0.005**	−1.7	0.10
T12 shear force	0.62 kN (3 ms)	1.5	0.08		
Shoulder-belt force	0.93 kN (20 ms)	2.8	**0.01**	0.1	0.94
HIC_{15}	203	0.4	0.69		
Shoulder-belt force	0.93 kN (20 ms)	2.8	**0.04**	0.73	0.60
N_{ij}	0.19	0.2	0.89		
Shoulder-belt force	0.93 kN (20 ms)	2.6	**0.01**	−0.85	0.61
Left femur force	2.1 kN (3 ms)	1.8	0.12		
Shoulder-belt force	0.93 kN (20 ms)	3.0	**0.005**	0.64	0.80
Right femur force	2.3 kN (3 ms)	1.0	0.40		
Shoulder-belt force	0.93 kN (20 ms)	2.7	**0.01**	−4.0	**0.05**
Femurs + lap belt	1.2 kN (20 ms)	1.9	0.10		

Note: main effect estimates are scaled to show the effect on R_{max} for changing each metric by the interquartile range (IQR). Interaction effect estimates are scaled to show the change in the IQR shoulder-belt effect on R_{max} with an IQR change in the second metric.

APPENDIX B:
FIELD INJURY MODELS

TABLE B.I
RESULTS OF LOGISTIC REGRESSION MODELS PREDICTING AIS≥2 THORACIC INJURY

		Baseline mean effects			Baseline Monte Carlo effects				
Metric	IQR	OR	95% CI	p-value	OR	95% CI	p-value	ACI	BSS
Bumper-firewall distance	12 cm	0.48	(0.25, 0.93)	**0.03**	0.51	(0.24, 1.03)	0.06	0.136	0.074
HIII NCAP acceleration	3.9 g (3 ms res.)	0.90	(0.43, 1.89)	0.77	0.79	(0.35, 1.74)	0.57	0.023	0.007
HIII NCAP deflection	6.5 mm	0.96	(0.43, 2.15)	0.91	1.10	(0.44, 2.74)	0.84	−0.008	0.001
HIII ODB acceleration	6.6 g (3 ms res.)	0.90	(0.40, 2.02)	0.80	1.07	(0.40, 2.86)	0.90	0.005	0.002
HIII ODB deflection	5.4 mm	1.87	(0.73, 4.78)	0.20	2.34	(0.83, 6.79)	0.11	0.079	0.046
Shoulder-belt force	0.69 kN (20 ms)	2.26	(1.09, 4.69)	**0.03**	2.55	(1.15, 6.05)	**0.03**	0.156	0.094
THOR femurs and lap belt	1 kN (20 ms)	0.62	(0.23, 1.65)	0.34	0.56	(0.18, 1.63)	0.29	0.045	0.028
THOR lower left deflection	4.6 mm	2.81	(1.22, 6.46)	**0.02**	2.74	(1.10, 7.15)	**0.03**	0.113	0.065
THOR maximum resultant deflection	4.2 mm	0.76	(0.47, 1.25)	0.29	0.83	(0.46, 1.48)	0.54	−0.012	0.022
THOR PC score	0.84	0.96	(0.48, 1.91)	0.91	1.04	(0.48, 2.18)	0.93	−0.001	0.000
THOR T12 shear force	0.62 kN (3 ms)	0.34	(0.15, 0.79)	**0.02**	0.34	(0.13, 0.81)	**0.02**	0.172	0.109
THOR T4 acceleration	6.6 g (3 ms res.)	1.61	(0.66, 3.90)	0.30	1.66	(0.58, 5.06)	0.36	−0.059	0.003
THOR T4 rotation	10 deg. (forward)	2.74	(1.25, 5.98)	**0.01**	2.75	(1.06, 7.60)	**0.04**	0.163	0.116
THOR thoracic injury criterion	21	0.69	(0.21, 2.28)	0.55	0.73	(0.19, 2.66)	0.64	−0.011	0.010

Note: to facilitate comparison between metrics, odds ratios (ORs) are scaled to show the effect of changing each metric by the interquartile range (IQR). For consistency with other metrics, HIII sternum deflection was inverted (compression positive). ACI = Average Classification Improvement (Eq. 1); BSS = Brier skill score (Eq. 3). Both ACI and BSS measure the degree of improved injury prediction relative to the baseline model using delta V (ΔV), age and weight without any crash test metrics. P−values ≤ 0.05 are in **bold**. res. = resultant.

TABLE B.II

RESULTS OF LOGISTIC REGRESSION MODELS PREDICTING AIS≥3 THORACIC INJURY

		Baseline mean effects			Baseline Monte Carlo effects				
Metric	IQR	OR	95% CI	p-value	OR	95% CI	p-value	ACI	BSS
Bumper-firewall distance	12 cm	0.63	(0.30, 1.33)	0.23	0.61	(0.27, 1.37)	0.24	0.118	0.004
HIII NCAP acceleration	3.9 g (3 ms res.)	0.73	(0.32, 1.69)	0.47	0.73	(0.31, 1.74)	0.48	0.006	0.008
HIII NCAP deflection	6.5 mm	0.87	(0.32, 2.33)	0.78	0.89	(0.33, 2.46)	0.83	0.006	0.005
HIII ODB acceleration	6.5 g (3 ms res.)	0.50	(0.19, 1.31)	0.17	0.49	(0.18, 1.34)	0.17	0.049	0.034
HIII ODB deflection	5.4 mm	0.75	(0.25, 2.26)	0.61	0.80	(0.26, 2.49)	0.70	0.014	0.000
Shoulder-belt force	0.67 kN (20 ms)	2.55	(1.06, 6.17)	**0.04**	2.68	(1.08, 6.94)	**0.04**	0.253	0.102
THOR femurs and lap belt	1 kN (20 ms)	1.66	(0.60, 4.62)	0.34	1.59	(0.55, 4.76)	0.40	0.021	0.002
THOR lower left deflection	4.6 mm	3.43	(1.32, 8.91)	**0.01**	3.53	(1.31, 9.90)	**0.01**	0.144	0.065
THOR maximum resultant deflection	4.2 mm	0.52	(0.29, 0.91)	**0.03**	0.52	(0.28, 0.96)	**0.04**	0.076	0.025
THOR PC score	0.84	0.76	(0.35, 1.67)	0.50	0.79	(0.34, 1.81)	0.59	−0.009	−0.001
THOR T12 shear force	0.62 kN (3 ms)	0.48	(0.19, 1.18)	0.12	0.46	(0.18, 1.18)	0.11	0.088	0.006
THOR T4 acceleration	6.6 g (3 ms res.)	1.03	(0.35, 2.98)	0.96	0.98	(0.28, 3.34)	0.98	−0.002	0.000
THOR T4 rotation	10 deg. (forward)	2.27	(0.95, 5.45)	0.07	2.22	(0.80, 6.02)	0.12	0.061	0.036
THOR thoracic injury criterion	21	0.29	(0.08, 1.11)	0.08	0.29	(0.07, 1.20)	0.09	0.029	0.000

Note: to facilitate comparison between metrics, odds ratios (ORs) are scaled to show effect of changing each metric by the interquartile range (IQR). For consistency with other metrics, HIII sternum deflection was inverted (compression positive). ACI = Average Classification Improvement (Eq. 1); BSS = Brier skill score (Eq. 3). Both ACI and BSS measure the degree of improved injury prediction relative to the baseline model using ΔV, age and weight without any crash test metrics. P-values ≤ 0.05 are in **bold**. res. = resultant.

APPENDIX C:
THOR CERTIFICATION FAILURES

Component test	Parameter	Peak value relative to calibration corridor
Face	Head resultant acceleration	+0.6%
Neck flexion	Upper neck Y moment	+1.0%
Neck flexion	Upper neck Z force	+0.4%
Neck extension	Upper neck Y moment	−4.1%
Neck extension	Upper neck Z force	−9.5%

APPENDIX D:
MONTE CARLO BASELINE INJURY RISK ESTIMATION

The logistic regression models of real-world injury outcomes based on test metrics included an offset for each of the 57 drivers to account for non-vehicle risks due to crash delta V, age and weight. These offsets were calculated using the effect estimates from "baseline" models of injury in the larger weighted sample of crashes. Figures 4 and 5 illustrate the results of the baseline models. To account for variability in the baseline model estimates, Monte Carlo predictions were performed using the estimated distributions for each baseline metric (delta V, age and weight). This process is described below, with examples given for the effects of delta V and R_{max} on AIS≥3 thoracic injury risk.

The results of the baseline AIS≥3 injury risk model, calculated using 902 NASS-CDS cases and their weights, are shown in Table D.I.

TABLE D.I
BASELINE AIS≥3 INJURY MODEL ESTIMATES

Term	Estimate	Standard error	p-value
Intercept	−13.810	2.293	<0.001
Delta V	0.113	0.016	<0.001
Nonlinear spline (age): 1	4.550	1.250	<0.001
Nonlinear spline (age): 2	3.060	1.794	0.09
Nonlinear spline (age): 3	5.497	1.055	<0.001
Weight	0.020	0.016	0.23

Driver age was modeled as a nonlinear predictor of injury. The estimates in Table D.I correspond to the individual components of the cubic spline function and are not interpretable on their own. Delta V and driver weight were modeled as linear predictors, and the estimates for these terms correspond to the slopes of the log-odds plots in Fig. 5. For each of the 57 drivers with representative THOR tests, the baseline injury log-odds y_i is calculated as:

$$y_i = \beta_{int} + \Delta V_i \times \beta_{\Delta V} + f(AGE|Age_i) + Weight_i \times \beta_{weight} \quad \text{(Eq. D.1)}$$

where β_{int}, $\beta_{\Delta V}$, and β_{weight} are the parameter estimates in Table D.I, $f(AGE)$ is the cubic spline function for age, and ΔV_i, Age_i and $Weight_i$ are the individual values for the case driver. The results of injury models including values of y as offsets are shown in Table B.II under "Baseline mean effects".

Model offsets calculated from the estimate point values in Table D.I do not account for the uncertainty associated with each estimate. The density function for the Gaussian distribution defined by the estimate and standard error for ΔV is shown in Fig. D.1. The Monte Carlo method was used to account for the uncertainty of baseline effects using the following procedure:

1. 1,000 samples were drawn randomly from each of the parameter estimate distributions for ΔV, age, weight, and the model intercept. Each sample represented an alternative estimate of baseline risk effects.
2. Each of the 1,000 samples was used to calculate new y values for the 57 cases according to Eq. D.1.
3. The effect of each test metric was modeled 1,000 times, with each model using one of the unique sets of y values as an offset.
4. The result of the modeling process was a set of 1,000 different estimated effects and standard errors for each test metric. Gaussian distributions representing each of these estimates were constructed. Fig. D.2 shows 10 of the 1,000 distributions for R_{max}.
5. The 1,000 distributions were combined by randomly drawing 1,000 values from each to produce a final sample of 1,000,000 effect estimates for each metric. The density function for the R_{max} sample is shown in Fig. D.3.
6. The mean of each test metric sample was taken as the final point estimate for that metric, the 95% confidence interval was defined by the 2.5 and 97.5 percentile values, and the z-score and p-value were calculated from the standard deviation and mean. Corresponding odds ratios scaled to the interquartile range (IQR) for each metric are shown in Table B.II under "Baseline Monte Carlo effects".

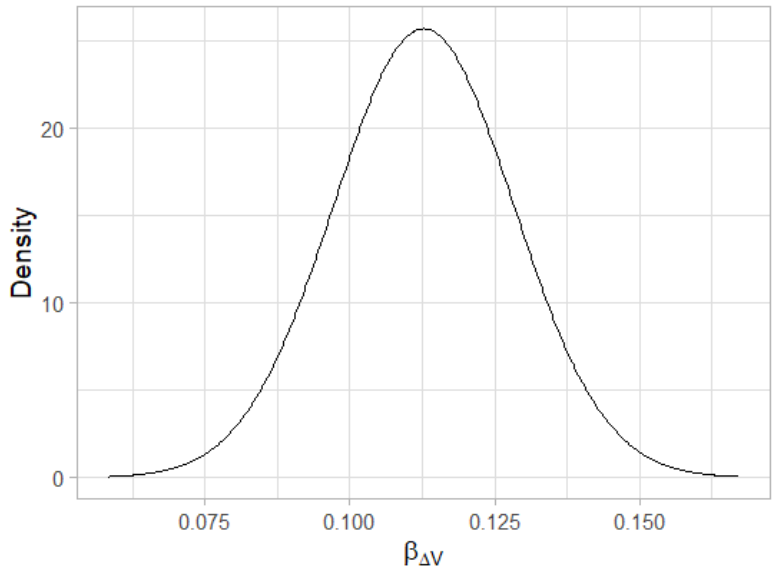

Figure D.1. Density function representing the estimated ΔV effect distribution.

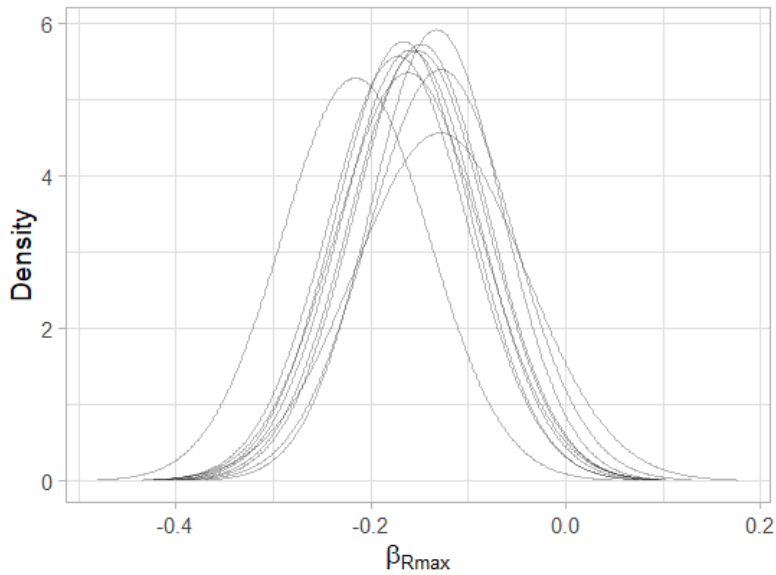

Fig. D.2. Density functions representing 10 of the 1,000 estimated R_{max} effect distributions.

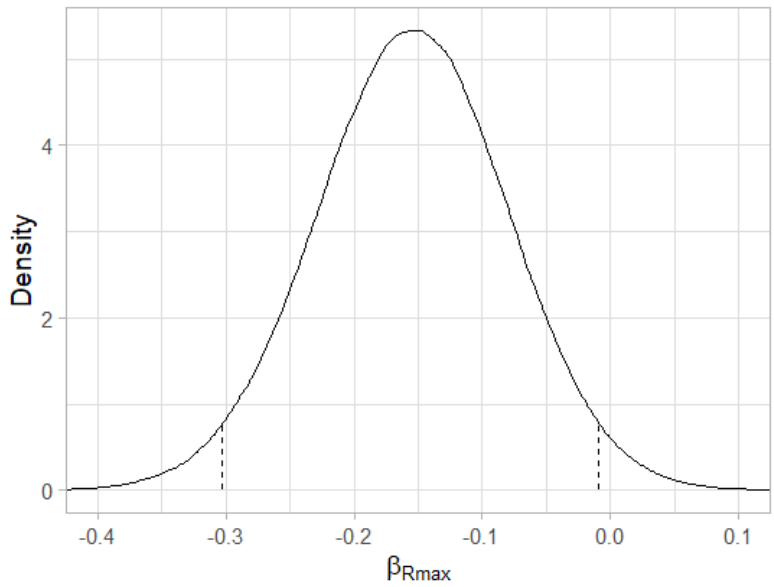

Fig. D.3. Final density function for the estimated R_{max} effect distribution. Dashed lines indicate the 2.5 and 97.5 percentile values.

APPENDIX E:
THOR CRASH TEST DETAILS

TABLE E.I
THOR CRASH TEST VEHICLES AND MATCHING FIELD CASES

Test ID	Test vehicle	Field cases
CF19020	2016 Toyota Camry	CDS: 2015-48-019-v1 CDS: 2015-49-109-v1 CISS: 6809-2
CF19021	2019 Toyota Camry	None
CF19024	2018 Volkswagen Atlas	None
CF19025	2019 Nissan Altima	None
CF19026	2018 Mazda 6	None
CF19027	2019 Chevrolet Equinox	CISS: 10616-2 CISS: 15525-2
CF19028	2017 Honda Civic	None
CF19029	2017 Chrysler Pacifica	None
CF19030	2020 Subaru Forester	None
CF19031	2020 Hyundai Santa Fe	None
CF19032	2019 Volvo XC60	None
CF19033	2020 Ford Escape	None
CF20001	2002 Ford Focus	CDS: 2009-02-155-v1 CDS: 2013-05-052-v1 CISS: 13562-1
CF20002	2008 Ford Fusion	CDS: 2009-76-128-v2 CDS: 2010-41-207-v2 CDS: 2013-73-033-v1
CF20003	2003 Mercedes-Benz ML 320	CDS: 2008-82-058-v1
CF20004	2007 Honda Accord	CDS: 2012-09-069-v1 CDS: 2012-49-063-v1 CDS: 2014-45-073-v2
CF20005	2013 Volvo XC90	CDS: 2012-45-087-v2
CF20006	2011 Honda Element	CDS: 2010-45-043-v1
CF20007	2007 Chevrolet Malibu	CDS: 2007-11-135-v1 CDS: 2013-13-048-v2
CF20008	2006 Chevrolet Uplander	CDS: 2012-02-059-v2 CDS: 2012-12-038-v2
CF20009	2016 Nissan Altima	CDS: 2013-49-134-v1 CISS: 15954-1
CF20010	2005 Buick LeSabre	CDS: 2006-48-085-v1 CDS: 2006-49-149-v3 CDS: 2009-12-136-v1 CDS: 2010-12-138-v1
CF20011	2004 Subaru Forester	CDS: 2008-02-112-v1
CF20012	2005 Toyota Camry	CDS: 2008-43-274-v2 CDS: 2008-45-170-v1 CDS: 2008-48-134-v2 CDS: 2008-79-091-v1 CDS: 2011-82-006-v1 CDS: 2012-05-016-v1 CDS: 2013-05-093-v1 CISS: 10982-2
CF20013	2007 Toyota Avalon	CDS: 2015-04-073-v2
CF20014	2007 Mazda CX-7	CDS: 2013-49-085-v1
CF20015	2005 Honda Civic	CDS: 2012-79-161-v1
CF20016	2011 Suzuki SX4	CDS: 2015-08-116-v2
CF20017	2011 Chevrolet Impala	CDS: 2013-74-097-v2 CDS: 2015-74-026-v2 CDS: 2015-81-051-v1
CF20018	2010 Nissan Altima	CDS: 2012-04-034-v1 CDS: 2013-74-118-v1 CISS: 15054-1
CF20019	2011 Honda Civic	CDS: 2008-43-212-v2 CDS: 2011-41-130-v2
CF21001	2001 Chevrolet Impala	CDS: 2007-12-200-v1 CDS: 2008-43-269-v2 CDS: 2009-73-054-v1
CF21002	2001 Toyota Highlander	CDS: 2009-09-189-v2 CDS: 2009-81-021-v1 CDS: 2012-43-102-v1 CDS: 2015-49-096-v2
CF21004	2009 Ford Focus	CDS: 2013-45-072-v1
CF21005	2009 Chevrolet Cobalt	CDS: 2008-73-093-v1
Note: tests without a matching field case did not contribute to the field injury models but were included in the correlation analysis (Appendix A)		

TABLE E.II
THOR ODB TEST SUMMARY MEASURES

	Resultant thoracic deflection						Belt force							Femur force	
Test ID	Upper right (mm)	Upper left (mm)	Lower right (mm)	Lower left (mm)	PC Score	Thoracic injury criterion	Shoulder (kN; 20 ms)	Lap (kN; 20 ms)	T4 thoracic spine accel. (g; 3 ms res.)	T4 thoracic spine rotation (deg.; forward)	T12 thoracic spine shear force (kN; 3 ms)	HIC_{15}	N_{ij}	Left (kN; 3 ms)	Right (kN; 3 ms)
CF19020	52.7	32.6	34.6	23.2	7.4	115	NA	2.28	32.4	50.3	−1.76	182	0.69	−2.58	−2.45
CF19021	47.9	35.6	31.5	15.8	6.7	101	NA	3.36	40.3	42.2	−1.47	194	0.49	−1.40	−1.53
CF19024	54.6	44.9	35.9	19.2	7.0	99	4.48	5.24	40.7	36.7	−2.25	383	0.61	−3.40	−1.49
CF19025	53.7	32.6	34.3	20.6	7.0	103	4.89	2.50	38.4	45.4	−1.97	161	0.60	−1.11	−1.51
CF19026	51.5	30.5	39.8	21.8	7.1	102	3.89	4.06	39.0	43.3	−1.69	212	0.53	−2.12	−1.63
CF19027	41.3	27.0	26.4	18.6	5.2	72	2.86	4.35	42.4	51.6	−2.41	180	0.52	−2.32	−3.94
CF19028	50.2	33.6	46.4	13.6	7.5	97	3.73	3.05	35.5	29.9	−0.94	240	0.47	−2.24	−2.64
CF19029	51.0	30.6	35.2	11.3	6.3	90	3.78	4.10	39.2	29.4	−1.47	207	0.68	−0.65	−1.68
CF19030	53.9	44.3	34.3	20.6	7.1	106	4.50	5.88	44.8	41.9	−2.41	255	0.71	−1.74	−0.99
CF19031	44.8	28.8	24.6	17.1	5.4	80	3.66	3.93	36.2	38.8	−2.18	333	0.45	−2.14	−1.62
CF19032	48.0	37.5	28.8	16.0	6.3	100	4.49	3.19	43.0	45.8	−1.64	279	0.66	−1.45	−1.46
CF19033	44.6	25.3	28.1	12.0	5.6	86	3.98	3.43	37.2	46.5	−1.39	215	0.53	−0.80	−1.20
CF20001	57.7	35.1	44.2	22.6	7.2	96	5.20	3.57	44.6	41.7	−1.79	268	0.70	−3.66	−3.09
CF20002	59.5	35.7	48.7	13.8	7.7	101	5.01	2.78	44.7	38.5	−1.34	210	0.67	−3.92	−3.98
CF20003	57.8	26.3	48.5	NA	NA	NA	NA	NA	41.7	54.0	−1.15	685	1.00	−5.04	−3.46
CF20004	53.2	32.5	43.2	16.4	7.2	104	5.20	4.65	38.1	28.0	−0.99	348	0.75	−1.16	−2.37
CF20005	46.0	25.2	35.1	19.5	6.1	83	NA	NA	46.2	54.7	−2.12	668	0.62	−4.89	−4.00
CF20006	49.4	34.1	36.8	19.8	7.1	111	NA	NA	35.7	43.7	−1.36	253	0.75	−4.58	−4.55
CF20007	53.5	29.0	39.9	15.4	7.0	103	4.91	2.63	36.4	49.5	−1.50	303	0.80	−3.59	−3.71
CF20008	53.5	30.3	37.2	16.9	6.8	102	3.75	1.76	43.0	30.2	−1.16	491	0.48	−5.10	−5.82
CF20009	50.4	30.0	30.6	18.4	6.1	92	4.06	4.26	44.5	35.5	−2.31	276	0.52	−0.22	−1.98
CF20010	55.9	27.2	42.0	13.9	7.4	112	4.75	0.76	48.3	39.2	−0.75	612	0.91	−4.76	−3.50
CF20011	55.6	34.5	45.9	10.1	7.2	91	4.66	2.57	35.1	48.1	−1.34	288	0.77	−4.02	−4.39
CF20012	51.8	28.0	39.4	13.5	6.6	91	4.38	3.00	41.5	35.6	−1.15	343	0.61	−2.41	−3.43
CF20013	52.3	28.8	33.4	17.7	6.4	103	4.66	4.76	42.5	39.9	−1.86	276	0.71	−1.36	−2.20
CF20014	49.6	NA	37.7	19.8	NA	NA	4.83	3.71	45.9	41.9	−2.54	429	0.57	−4.69	−2.11
CF20015	59.3	34.1	39.6	21.0	7.7	126	6.18	3.66	48.0	28.8	−1.28	582	0.90	−4.81	−4.09
CF20016	52.1	33.1	45.5	20.7	6.9	84	4.17	5.13	48.0	45.4	−2.81	540	0.59	−3.86	−4.15
CF20017	50.5	28.8	29.1	15.1	5.8	89	3.12	3.56	49.3	35.0	−1.80	425	0.59	−4.12	−4.80
CF20018	52.4	NA	37.1	11.5	NA	NA	4.30	4.05	36.2	25.6	−1.72	126	0.60	−2.52	−1.78
CF20019	54.3	28.8	54.5	14.6	8.2	113	NA	3.39	44.9	25.5	−1.31	641	0.82	−3.93	−2.98
CF21001	46.6	29.3	23.1	16.1	5.5	95	4.09	1.03	37.0	31.3	−0.83	500	0.48	−4.60	−7.81
CF21002	66.0	38.1	50.2	15.1	8.2	113	4.46	3.71	39.6	40.1	−1.50	411	0.45	−3.76	−3.55
CF21004	56.1	29.0	44.8	7.8	7.5	111	2.64	5.76	32.7	19.2	−0.46	233	0.49	−3.10	−3.02
CF21005	55.6	33.1	41.9	13.8	7.2	100	3.81	3.36	38.7	26.1	−1.62	296	0.38	−3.52	−5.01

TABLE E.III
THOR ODB TEST CLEARANCE MEASURES

Test ID	SHH (mm)	SHV (mm)	PA (deg.)	Head angle (deg.)	TSA (deg.)	NR (mm)	SCR (mm)	KDL (mm)	KDR (mm)
CF19020	−214	−226	34.0	−5.4	21.5	530	420	218	210
CF19021	−213	−212	35.0	−10.3	18.0	468	383	227	232
CF19024	−172	−234	33.7	−3.8	20.9	557	484	244	232
CF19025	−188	−178	31.3	−7.9	20.2	495	411	191	185
CF19026	−151	−199	30.5	−5.3	21.4	500	410	200	195
CF19027	−132	−112	32.7	−6.2	20.4	530	413	190	181
CF19028	−178	−256	32.9	−7.5	20.2	494	422	251	231
CF19029	−195	−83	30.7	−2.0	23.5	538	410	220	208
CF19030	−205	−194	35.2	NA	21.5	520	398	183	175
CF19031	−155	−251	34.6	−6.1	19.6	520	400	181	178
CF19032	−117	−120	32.2	−3.3	22.4	540	418	252	241
CF19033	−141	−100	33.5	−4.5	21.3	527	410	244	225
CF20001	−140	−136	33.3	−3.6	20.4	535	460	185	150
CF20002	−155	−126	31.1	−6.3	18.9	500	430	200	175
CF20003	−188	−127	30.6	1.8	23.6	569	450	130	110
CF20004	−171	−115	30.8	−2.0	21.6	514	420	227	211
CF20005	−186	−78	32.1	−0.8	23.6	557	430	180	174
CF20006	−218	−133	30.5	−4.0	21.5	561	460	142	144
CF20007	−186	−143	30.5	−0.2	22.2	530	470	185	150
CF20008	−197	−104	33.6	−3.3	19.2	500	410	208	200
CF20009	−195	−140	31.3	−4.7	21.7	505	400	216	212
CF20010	−184	−102	32.4	2.0	24.5	535	432	168	155
CF20011	−139	−66	30.3	−4.2	20.6	510	417	170	111
CF20012	−171	−203	35.3	−4.4	19.9	534	440	233	182
CF20013	−208	−201	30.6	−3.0	20.1	549	415	181	167
CF20014	−156	−97	32.2	0.4	22.8	570	452	231	213
CF20015	−152	−135	34.9	−3.0	20.1	530	450	230	215
CF20016	−114	−168	33.4	−5.1	20.3	560	454	227	208
CF20017	−182	−81	34.2	−1.8	21.5	524	440	255	237
CF20018	−188	−121	30.9	−9.2	14.5	445	400	220	185
CF20019	−196	−157	32.9	−6.1	19.8	502	424	225	242
CF21001	−174	−96	34.6	−2.4	23.8	534	442	215	231
CF21002	−193	−111	33.9	−6.0	20.0	546	433	180	198
CF21004	−401	−130	31.7	−11.0	15.5	470	450	220	205
CF21005	−473	−101	33.9	−3.0	21.7	538	455	214	223

Note: SHH = striker to H-point, horizontal; SHV = striker to H-point, vertical; PA = pelvic angle; TSA = thoracic spine angle; NR = nose to rim; SCR = steering wheel to chest reference; KDL = knee to dash, left; KDR = knee to dash, right.