Dynamic Research, Inc.

COMMENTS ON MONASH/ISCRR REPORT BY WORDLEY (2012) ENTITLED
"QUAD BIKE CRUSH PROTECTION DEVICES (CPDs): UPDATES TO ISCRR SNAPSHOT
REVIEW C-I-12-022"

DRI-TR-14-12

J. W. Zellner

R. M. Van Auken

S. A. Kebschull

T. A. Smith

22 November 2014

355 Van Ness Ave • Torrance • California 90501 • 310-212-5211 • Fax 310-212-5046 • www.dynres.com

TABLE OF CONTENTS

EXECUTIVE SUMMARY

EXECUTIVE SUMMARY

This report provides comments on a report by Wordley (2012) entitled "Quad bike Crush Protection Devices (CPDs): Updates to ISCRR Snapshot Review C-I-12-022". This Wordley (2012) report supplements the earlier Wordley and Field (2012) report and is primarily a commentary upon two additional publications by Dynamic Research, Inc. (DRI) relating to the effectiveness of an Australian designed and manufactured CPD, the Quadbar^{TM}, which is proposed to be fitted to All Terrain Vehicles (ATVs). Those DRI publications are Zellner et al. (2012), "Updated Injury Risk/Benefit Analysis of Quadbar Crush Protection Device (CPD) for All-Terrain Vehicles (ATVs)" and Zellner, et al. (2012a), "Replies to Lower (2011) Comments".

Regrettably, most of the commentary in the Wordley (2012) report comprises interpretations and analyses of, in particular, the Zellner, et al. (2012) report, which are frequently inaccurate, somewhat misleading, based on false premises or misunderstandings of relevant technical issues, without basis or incomplete. In addition, many of the comments in the Wordley (2012) report are now out of date as a result of more recent DRI research (i.e., Zellner, et al. (2014a) and Zellner, et al. (2014b)).

The principal deficiencies in the Wordley (2012) report include that:

– Wordley (2012) bases many of its conclusions about the Quadbar entirely upon the invalid test methods of Snook (2009), in which an ATV was tested alone, without any rider surrogate. Consequently, those tests took no account of where a rider would be or how, and at what point, he would separate from the ATV in the event of an overturn but, rather, made baseless assumptions as to where a hypothetical rider "would be", with inexplicably different assumptions as for the baseline ATV and the Quadbar ATV vehicle configurations, including that a rider would *always* come to rest under an upside down (non-Quadbar) ATV.

  Snook (2009) also failed to make any direct measurement of injury assessment variables. Instead Snook (2009) assumed, incorrectly, arbitrarily and without any indicated basis, that some particular outcomes would be injurious, or more injurious than other outcomes. For example, it was assumed that an ATV coming to rest upside-down was injurious, whereas an ATV coming to rest on its side was mistakenly[1] assumed to be harmless, and that an ATV continuing to roll "farther"

[1] For example, "A Handbook for Workplaces: Quad Bikes on Farms" (Work Safe Victoria, Edition No. 2, August 2009, p6) cites a case of "A 75-year-old man was killed whilst operating a quad bike equipped with a 50-litre spray tank full of chemical spray. He was working in a wet area on an incline of 20–30 degrees. The man's wife discovered him lying face down towards the rear of the quad bike which was on its side on top of the man." Another example was a New Zealand ATV on-its-side case described at the Technical Engineering Group meetings, Sydney, 5 to 7 October 2010, resulting in long term rider entrapment and subsequent medical amputation of the leg. The Peter Crole ATV fatal accident in Australia involved an ATV coming to rest on its side on the rider, with heavy equipment on front and rear cargo racks

(i.e., farther away from a rider who as a result of natural forces typically separates from the ATV at an early stage of an overturn) was more injurious than an ATV coming to rest in closer proximity to the rider, which are highly questionable assumptions for which there is contradictory evidence.

Despite the many serious flaws in the Snook (2009) research, Wordley (2012) states "***The limited number of experimental and simulation results which were considered valid [and primary among them, Snook (2009)] indicated that CPDs demonstrated the potential to reduce rider harm in low speed roll-over events." (p1, paragraph 2).***

– Wordley (2012) wrongly asserts that there is "***A significant difference between the style/standard of helmet being simulated (full-face) and that which is currently recommended by the industry and government bodies for Quad bike use (i.e. compliant with AS/NZS:1698)" (p1, 3rd bullet).***

  In fact:
  - the full-face (Bieffe B12) helmet simulated in Zellner et al. (2012) is *fully compliant* with the impact and coverage requirements of AS/NZS:1698-2006 (as well as ECE R22.[2]);
  - the ATV manufacturers' recommendations regarding helmets *do not prescribe a particular style/type of helmet* but, rather, that they comply with applicable federal, state and/or other standards[3] and that they are consistent with the guidelines provided by the ATV Safety Institute;

  In any event, the THH T70 half helmet suggested by Wordley (2012) was used by Zellner et al (2014b) in their extended simulation analysis, and this resulted in no change to the overall conclusions about the Quadbar in comparison to the Zellner et al. (2012) research.

– Contrary to the statement in Wordley (2012) that "***Predicted Injury/Benefit ratios from this updated work have been presented using a "single baseline" method rather than the commonly accepted and more logical "multiple baseline" method." (p1, 3rd bullet),*** in fact both the "single baseline" results and the "two

---

(Rechnitzer and Grzebieta, 2004). In Zellner et al. (2004), Appendix E, two of the three UK/US asphyxia cases involved unknown resting orientations, and the third involved wheel/face contact (which cannot be an upside-down orientation). In his findings in the inquest into the death of H, a child, Queensland Deputy State Coroner Lock was unable to ascertain whether an ATV had rolled ¼ times or 1¼ times onto its side (although he found it "difficult to conceive the latter figure"), but either way, the child had died of asphyxia after being pinned beneath the side of the ATV. In addition, Figures 10 and 11 in Zellner et al. (2012) indicate that based on 3,080 overturn simulations, the baseline ATV came to rest more frequently on its side, and that adding a Quadbar more than doubled the number of pinned/on-its-side/predicted asphyxiations.

[2]Smith, T.A, Koerner, J.E., <u>Certification Test Report according to AS NZS:1698-2006</u>, DRI-TM-14-67, 16 September 2014.

[3]For example, in the US, US/DOT/FMVSS 218 (which is very similar to AS/NZS:1698) which likewise does not specify the style of helmet.

baseline" results were presented in Zellner, et al. (2012), in Section IV.B and Appendix I, respectively. This comment is also now out of date as both the "single baseline" results and the "two baseline" results were presented *within the same portion of the text*, in the most recent Zellner, et al. (2014b) published paper. In addition, this statement is inaccurate as the "two baseline" method is not more "commonly accepted" compared to the "single baseline" method: there are no standards requiring it, and, far from being "commonly accepted", there is in fact no other original research in this field, other than DRI's, that uses a "two baseline" method.

– Despite having previously (in Wordley and Field (2012)) rejected the *specific simulation sample* used by DRI, Wordley (2012) relies upon that sample to inaccurately and misleadingly claim that "***according to DRI's updated research, the use of a CPD actually returns a risk/benefit ratio of 68% [42%, 114%] (slightly safer) for the unhelmeted condition, and 108% [69%, 169%] (marginally less safe) for the helmeted condition." (p2, paragraph 2)***.

  This claim ignores the fact that the results for the *population* of overturns reported in Zellner et al. (2012) indicate that the effects of the Quadbar are, *for both the helmeted and the unhelmeted condition, statistically insignificant* within a 95% confidence interval, the latter confidence interval (equivalent to p=0.05, Box, et al. (1978)) being by far the most widely accepted standard for being "somewhat convinced" of the reality of an outcome. The fact that the square bracketed percentages span the 100% value indicates that the 68% and 108% ratios (the point estimates for the simulation sample only) are, for the population of all overturns, *statistically indiscernible* from 100%, i.e., that they are in fact *statistically insignificant*. They are categorically *not* "slightly safer" or "marginally less safe", which descriptions would *only* apply to the specific UK/US simulation sample that Wordley and Field (2012) rejected as being invalid.

– The statement by Wordley (2012) that ***"for the unhelmeted condition the CPD was very close (within 14%) to returning a statistically significant prediction of increased safety."(p 25, 2nd bullet)*** is contrary to standard practice in statistical analysis, which does not involve stating that a factor is "very close" to a statistically significant prediction. The statistical test in and of itself returns either a statistically significant or a statistically not significant outcome. The power of the statistical test confirms the reliability of the outcome, which in this case was a non-statistically significant difference between the risk/benefit ratio with a CPD and without a CPD, for both helmeted and unhelmeted conditions.

  Moreover, the fact that the probability value (p=0.153) for the Quadbar unhelmeted risk/benefit outcome is *more than three times* the "conventional 'critical' significance level in common use" (i.e., p=0.05, Box, et al (1978)) indicates that the Quadbar was *not* "very close" to returning a statistically significant prediction of increased safety.

– The suggestion by Wordley (2012) that ***"A preliminary standard be proposed for the design and specification of Quad bike CPDs, perhaps initially based upon those for Tractor ROPS, or the performance of the CPD when subjected to the experiments proposed below." (p 26, paragraph 1)*** incorrectly pre-supposes that:
    - some scientifically valid or comprehensive data currently exist that indicate that a CPD on an ATV has more injury benefit than injury risk, across a realistic range of overturn conditions, when in fact there are none;
    - tractor ROPS are in some respect comparable with or relevant to the design of a CPD for an ATV, whereas:
        (a) under Australian state law, ROPS are not required for tractors of less than 560 kg mass[4], while a typical ATV has a mass of only 275 kg; and
        (b) tractors are totally different types of vehicles from ATVs, as tractors with ROPS are not "rider active" vehicles, have much greater size and mass, may have large unguarded wheels in close proximity to the driver, typically have seat belts and have an entirely different type of occupant seat, and have very different rollover motions and occupant motions.

– In stating that ***"The overall validity and usefulness of a correlation resulting from such a heavily filtered [accident] dataset was considered minor. The potential for researchers [to] tune their AIS coding of injuries to match their simulated results was also noted" (p27, paragraph 5)***, Wordley (2012) misconceives the purpose of the injury coding in the subject research, which was *to enable calibration of the simulation model.* This "filtering" is necessary and entirely appropriate because, at the current state of technology, simulation models are able to predict only a limited set of locations, types and severities of injury (this is the case for all crash dummies worldwide, including those applicable to cars, motorcycles, etc.). It would therefore be *inappropriate, senseless and impossible to attempt to include and calibrate those injuries that cannot be monitored by a crash dummy, which is what Wordley (2012) appears to recommend doing*. To include injuries that cannot be experimentally monitored and verified would be to attempt to compare "apples to oranges". This is the appropriate, and only, reason that some locations, types and severities of injuries were "filtered" (i.e., removed) from consideration.

  It is to be regretted that the failure of Wordley (2012) to appreciate the previous point has apparently prompted the accusation of impropriety implied in the statement, ***"the potential for researchers to tune their AIS codings…was noted***". This accusation is baseless and defamatory and is refuted.

[4]For example, Chapter 5, Part 5.1, Division 7, Subdivision 3, Clause 216 of the current NSW Work Health and Safety Regulation 2011. Also Australian Standard AS 1631.1-1996.

– The recommendation by Wordley (2012) that ***"In the mean time [sic], regulatory bodies should consider recommending the use of appropriately tested crush protection devices (CPDs) for riders who use Quad bikes at low speeds in the workplace and on farms." (p27)***, is without basis and inappropriate because:
    - the highly deficient Snook (2009) tests did not indicate any ***"potential to reduce harm"*** in any scientifically valid way;
    - the updated Zellner et al. (2012) research and overall findings, which addressed all of the supposed ***"serious limitations"*** previously suggested by Wordley and Field (2012), found that the Quadbar CPD has *statistically insignificant effects* on injuries and fatalities (as well as a risk/benefit percentage that is larger than that for any known vehicle safety device); and
    - there are no scientifically valid data which indicate that CPDs in general, or the Quadbar in particular, are more effective and less harmful at low speeds than at any other speeds.

Absent, therefore, any evidence to support the safety or efficacy of the Quadbar, the recommendation that they be fitted nonetheless is of doubtful ethical standing.

SECTION I

INTRODUCTION

This report provides comments on a report by Wordley (2012) entitled “Quad bike Crush Protection Devices (CPDs): Updates to ISCRR Snapshot Review C-I-12-022”. The Wordley (2012) report is an updated commentary supplementing its predecessor, the Wordley and Field (2012) report. Wordley (2012) primarily addresses two more recent publications by Dynamic Research, Inc. (DRI) relating to the effectiveness of an Australian designed and manufactured CPD, referred to as the Quadbar™. For the most part, the Wordley (2012) update focuses on a report by Zellner, et al. (2012) entitled “Updated Injury Risk/Benefit Analysis of Quadbar Crush Protection Device (CPD) for All-Terrain Vehicles (ATVs)”, and to a lesser extent, on a report by Zellner, et al. (2012a) entitled “Replies to Lower (2011) Comments”.

In the current report, the term “All Terrain Vehicle” refers to a “A motorized off-highway vehicle designed to travel on four low pressure tires, having a seat designed to be straddled by the operator and handlebars for steering control” as defined in ANSI-SVIA 1-2007, which in April 2009 became a mandatory safety standard in the United States, which accounts for more than 90% of ATV production and sales worldwide. In this report, the term “quad bike” (or “quad”) is not used (other than where it forms part of a quoted passage) because it is a slang, non-legally precise term used in some regions, which may include ATVs, but which has also been used to refer to other, diverse, four wheeled vehicles that do not comply with ANSI/SVIA-1-2010. The term "CPD" refers to “Crush Protection Devices”, such as the Quadbar, that have a structure intended to provide overturn protection and that do not have rider restraint belts. The term "ROPS" refers to “Roll Over Protection Systems” that have a structure intended to provide overturn protection and that have one or more restraints (e.g., belts, seat back, etc.). The term "overturn" is considered to include rollovers, pitchovers and combined axis overturns of a vehicle. The term "misuse" refers to using a product in a way that is contrary to the warning labels that are required on ANSI/SVIA 1-2010-compliant ATVs, for example, not wearing a helmet, use of adult-sized ATVs by children, etc.

The Wordley (2012) report did not take into account, and the report does not reference, the Zellner, et al. (2012b) report entitled “Comments on Monash/ISCRR Report by Wordley and Field (2012) Entitled “Quadbike Safety Devices: A Snapshot Review”. So, unfortunately, Wordley (2012) continues to repeat points made and already responded to in the latter report, which was available on the internet as of 17 October 2012 (and of which ISCRR was advised on that same date). This was several weeks before the finalization of the Wordley (2012) report and several months before the Wordley (2012) report became generally available.

Subsequent to the Wordley (2012) report, two subsequent DRI papers were published: Zellner, et al. (2014a) and Zellner, et al. (2014b), each describing further simulation results and the latter paper also addressing several issues raised by Wordley

(2012). Wordley (2012) does not, of course, take those papers into account and it must therefore be read subject to them.

This report proceeds by presenting the respective quotation from the Wordley (2012) text in ***bold italics***, including page and paragraph, which is followed by a comment by the authors of the current report. For clarification purposes, insertions to quotations are included in square brackets [ ] as to source, topic and so forth. Note that the quotations are organized according to the section and subsection in which they appear in the Wordley (2012) report.

Note that the Wordley (2012) report is repetitious, involving recapitulation of the same or similar discussions in several different sections, and restatements (often with different emphases or points) in the overviews and summaries. Therefore, for the sake of completeness, but also in order to minimize repetition of comments, for a given quotation, each comment is intended either to be self-contained, or to be in summary form, cross-referencing other closely related comments in the current report. The comments are followed by the summary and conclusions of the current report in Section VIII, and a list of cited references.

The comments in the current report include comments on the information provided in the Wordley (2012) report, as well as, in some cases, information contained in the original references reviewed in the Wordley (2012) report.

The authors of this report have conducted extensive original research on the subject of ATV ROPS/CPDs beginning in 1996, and a small amount of their work is discussed in Wordley (2012). The authors have extensive experience in operating and testing four wheeled ATVs since their invention in 1982, and modelling four wheeled ATVs since 1988. They have been extensively involved in vehicle crash protection and crash avoidance studies involving dozens of vehicle categories in hundreds of projects for the US Government as well as for vehicle manufacturers and manufacturer associations worldwide. These have included investigations of, amongst other things, evaluation of roll over protection systems for a wide variety of vehicles; and investigating roll bar decapitations involving overturns of off-road open vehicles (e.g., dune buggy; jeep) where the involved drivers did not use the provided restraint belts.

The Wordley (2012) report is stated to have ***"been overseen by a committee comprising… [amongst others] Yossi Berger (AWU), and Tony Lower (University of Sydney)"*** (p6, paragraph 4)). It is noted that Dr. Berger's organization and Dr. Lower were both outspoken advocates of fitting CPDs to ATVs prior to publication of both this Wordley (2012) report and the prior report, Wordley and Field (2012). The report does not describe what steps were taken to ensure that this prior advocacy did not influence the numerous opinions expressed in the report in support of the Quadbar and other CPDs.

SECTION II

COMMENTS ON "EXECUTIVE SUMMARY"

The comments on the Wordley (2012) Executive Summary are as follows:

1. ***"[In the initial Wordley and Field (2012) review, a] number of serious limitations were identified with respect to the scenarios chosen and simulation methods utilised by the existing research." (p1, paragraph 2)***

   To the extent that this statement refers to DRI research, all of the so-called "serious limitations" have either been fully explained (e.g., as being due to misunderstandings by the critics) or, if arguably valid, were fully accounted for in the Zellner, et al. (2012) research. This point was also fully explained in Zellner, et al. (2012b), which was accessible on the internet[5] as of 17 October, 2012 (of which ISCRR was advised on the same date), several weeks before the Wordley (2012) report was finalized, and several months before the Wordley (2012) report became generally available.

2. ***"The limited number of experimental and simulation results which were considered valid, indicated that CPDs demonstrated the potential to reduce rider harm in low speed roll-over events." (p1, paragraph 2)***

   First, it is unclear exactly to which studies this refers, and by whom these experimental and simulation results "were considered valid". To be meaningful and traceable, this statement should say something like "results of Snook (2009)…which were considered valid *by the authors of Wordley and Field (2012)*".

   Second, the statement is inaccurate. Describing the results as "valid" ignores the 20 or more substantial issues with the validity of the associated Snook (2009) tests, as described in Zellner, et al. (2012b), and as summarized in the current report.

   Third, the statement, "CPDs demonstrated potential to reduce harm" is inaccurate, as the "demonstration" refers to the invalid test methods of Snook (2009), which:
   - tested an ATV-alone, without any rider surrogate;
   - failed to account for where a rider would be or how and when he would separate from the ATV in the event of an overturn; and
   - assumed incorrectly and without any basis that some outcomes would be injurious (e.g., the ATV landing on its top at a given point on the ground, or continuing to roll further beyond this, whereas, by this point many or most riders would actually have separated from the ATV) and that other outcomes would *not* be injurious (e.g., the Quadbar ATV coming to rest on its side or on its wheels, whereas this is known to occur among fatal

[5] Accessed at www.dri-atv-rops-research.com, 17 October, 2012.

entrapment/asphyxiation cases[6], or the ATV *not* continuing to roll further away from the rider).

3. ***"The initial review suggested that a range of further (predominantly experimental) research be conducted to confirm and quantify these predicted benefits [in low speed roll-over events]." (p1, paragraph 2)***

   This statement (in common with the associated statements in the original Wordley and Field (2012) report) is unclear, as it is not discussed how a sufficient number of tests can be conducted that can adequately cover the full range of real ATV usage (and overturn) conditions. Focusing on only a few conditions means that *in other conditions* it is unknown whether the CPD is beneficial, harmful or has no effect. In addition, as discussed immediately above, this statement is based on a false premise that the Snook (2009) "Vehicle Accelerator" tests are valid, as that test methodology was based on unrealistic ATV motions (e.g., sliding sideways down a slope) and ignored any form of injury monitoring using, e.g., an injury monitoring crash dummy and objectively measured injury assessment variables. Furthermore, the absence of a rider surrogate in the Snook (2009) tests, the mass of which is a substantial fraction of the ATV mass, would be expected to have a large influence on the overturn motions of the ATV (at least until ATV/rider separation). Addressing these issues is the reason why computer simulations with a crash dummy, across a wide range of accident-based conditions (as well as a series of actual full-scale tests done with a dummy on natural terrain, in order to calibrate/validate the simulations), were used by DRI to evaluate proposed ATV CPDs and ROPS.

4. ***"In the meantime, the use of appropriately tested CPDs was recommended for riders who use Quad bikes at low speeds on farms and in other work places." (p1, paragraph 2)***

   First, this statement contains an ethical incongruity in, on the one hand, recommending tests to "confirm" the existence of a benefit, while also recommending the use of the relevant device in the meantime when, by definition, the benefit (not to mention any unintended risks) has yet to be confirmed.

   Second, this statement is based on the false premise that the highly deficient Snook (2009) tests indicated "potential to reduce harm" (when in fact they did not do so: arguably, they showed no change in or increased potential for harm, as noted above). In addition, this statement is without clear basis in terms of the "low speed" aspect. There are no scientifically valid data which show that CPDs in general, or the Quadbar in particular, are more effective and less harmful at low speeds than at any other speeds. In addition, Wordley does not address potential harm of CPD's at higher speeds (i.e., greater than the maximum speed of 11 km/h used in the Snook (2009) tests), at which ATVs frequently travel. None of the Snook tests considered such conditions. The Zellner et al (2012) research has found that a device such as a Quadbar can be injurious to the rider at speeds below, at and above 11 km/h. In any

[6] See Footnote 1.

case, even if the harmful effects of the Quadbar were limited to above 11 km/h (which they are not), there is no practical way to ensure that the ATV/rider does not overturn at speeds above 11 km/h, thereby potentially encountering such harmful effects of the Quadbar.

5. ***"A number of detail changes have been made to the simulation model to address some of the limitations identified in the initial work." (p 1, 1st bullet);***

This statement - particularly the word "some" – implies that there were other limitations in the initial work which have not yet been addressed. This would not be accurate, as explained in the response in Section II.1 above and in Zellner, et al. (2012b), because all of the claimed "limitations" were either fully explained (e.g., as being due to misunderstandings of, or false premises assumed by, the critics), or if arguably valid, were fully accounted for in the Zellner, et al. (2012) research by means of the "detail changes" noted in this statement.

6. ***"A significant difference between the style/standard of helmet being simulated (full-face) and that which is currently recommended by the industry and government bodies for Quad bike use (i.e. compliant with AS/NZS:1698)." (p1, 3rd bullet);***

This statement is inaccurate and represents a major misunderstanding in Wordley (2012). First, the full-face Bieffe B12 helmet used in the Zellner et al. (2012) and earlier DRI research is fully compliant with the impact and coverage requirements of AS/NZS:1698.[7] Second, AS/NZS:1698-2006 does not prescribe any particular style/type of helmet. Both full-face and open-face motorcycle helmets that are universally used by motorcyclists throughout Australia comply with AS/NZS:1698, as do some half-helmets (also known as "shorty" helmets) such as the THH T70 suggested by Wordley (2012). Third, the ATV manufacturers' recommendations regarding helmets are that they comply with the applicable federal, state and/or other standards[8] and are consistent with the guidelines provided by the ATV Safety Institute as follows:

> "Select a helmet that meets or exceeds your state's safety standards and carries either the Department of Transportation (DOT) label or the Snell Memorial Foundation label. Your helmet should fit snugly and fasten securely. Full-face helmets help protect your face as well as your head. Open-face helmets are lighter and may be cooler, but should be used with mouth protection. Eye protection should be used with both types of helmets."[9]

---

[7]See Footnote 2.
[8]See Footnote 3.
[9]Anon., "Tips and Practice Guide for the All-Terrain Vehicle Rider", page 4, www.atvsafety.org/InfoSheets/ATV_Riding_Tips.pdf, ATV Safety Institute, a Division of the Specialty Vehicle Institute of America, (accessed 9 July 2013).

Note that the only ATV helmets currently illustrated in the aforementioned guideline are *full-face* helmets. Also note that inquiries indicate that neither the FCAI nor the largest Australian ATV importer specifically recommend "shorty" (i.e., half) helmets.

In any event, the speculation inherent in this statement - that significantly different outcomes would result if a different helmet was simulated – is without basis. As further discussed in Section VI.3 and in Zellner, et al. (2014b), substituting the half-helmet THH T70 characteristics into the DRI simulation model and re-running the simulations and risk-benefit analysis indicates that the Quadbar risk/benefit and net benefit outcomes are <u>similar</u> to those with the full-face (Bieffe B12) helmet, and do not change any of the conclusions regarding the Quadbar. As might be expected, the net benefit of the half-helmet alone (as a protective device) is less than that of the full-face helmet alone, but still substantial. Note that Rodgers (1990) of the US/CPSC reported net benefit of 64% for ATV helmets for non-fatal head injuries, which is very similar to the net benefits predicted for full-face helmets by the DRI simulations (i.e., 60% net benefit for predominantly non-fatal injuries as reported in Zellner et al. (2014b)).

7. ***"Major changes were made to the recorded injury data set which was used to compare with the simulated injuries." (p1, 3rd bullet);***

   This statement is inaccurate, as the "recorded injury data set" was not changed but, rather, the injury codings were updated and revised to include only those injuries that are potentially related to dummy objective injury measurements. These changes to the injury codings were required and appropriate because the specific purpose of the subject research was to enable calibration[10] of the outcomes of the simulation model against the outcomes of real accidents. At the current state of technology, simulation models are only able to predict specific locations, types and severities of injury (as is the case for all crash dummies used in crash tests with cars, motorcycles, etc.).

   Contrary to what is mistakenly implied by Wordley (2012), the range of injury prediction that is feasible with crash dummies could not be validly extended without, in general, years of research and often extensive peer review by standards and/or scientific conference committees. This is further discussed in SectionV.24.

8. ***"Predicted Injury/Benefit ratios from this updated work have been presented using a "single baseline" method rather than the commonly accepted and more logical "multiple baseline" method." (p1, 3rd bullet).***

   This statement is inaccurate and could easily give rise to misunderstanding. In fact, both the "single baseline" results and the "two baseline" results were presented in Section IV.B and in Appendix I, respectively, of Zellner, et al. (2012). This is also true of the later published papers, i.e., Zellner et al. (2014a) (in which single baseline results are presented in Appendix B and two baseline results are presented in

[10]This calibration was an aggregated frequency distribution level, as discussed in Zellner, et al. (2012), Appendix G, and Zellner et al. (2012a).

Appendix C) and Zellner, et al. (2014b) (in which both single baseline and two baseline results are presented in the main text). The main text discussion of Zellner, et al. (2012), which is based on the "single baseline" method, cross-references (on page 64) the Appendix I "two baseline discussion".

In addition, this statement is inaccurate as the "two baseline" method is *not* more "commonly accepted" compared to the "single baseline" method; there are no standards requiring it, and in fact there is no other original research in this field, other than DRI's, that uses a "two baseline" method. All of the research that uses the ISO 13232 methodology of which the authors are aware used the "single baseline" methodology, and the ISO 13232 standard itself only references a single baseline configuration.

The single baseline method is relevant because (a) it is relative to the intended (and safer) condition (i.e., helmeted) and (b) importantly, it can capture what occurs if a user were to mistakenly assume that the CPD has a significant protective effect and that he/she no longer needs to wear a helmet. As further discussed in Section V.36 herein, the "two-baseline" approach can be misleading in that it shifts the focus away from a primary hazard, which in this case is not wearing a helmet. Not wearing a helmet is approximately 2.5 more hazardous than wearing a helmet, and even if a device such as a CPD reduced that hazard by a small but statistically significant amount – which the DRI data *do not indicate* – the unhelmeted hazard – and the large benefits of wearing a helmet – far exceed the estimated effects of fitting a CPD while unhelmeted, which were found by the DRI research to be statistically insignificant.

In addition, the validity of the statement that the two baseline method is "more logical" depends on the reader's perspective. If the reader is an ATV safety advocacy group or ATV manufacturer, which is recommending that all riders wear helmets, then a single baseline, the intended "helmeted" condition, is more logical.

9. ***"A risk/benefit ratio of 492% is quoted (much less safe), in comparison to the 71% (slightly safer) returned by their previous study of the same device. Detail analysis found that this increase could be attributed to their adoption of a "single baseline" comparison method. This reported increase was implied by calculating the injury/benefit ratios for an unhelmeted rider with a CPD, via comparison with a helmeted rider, without the CPD. Such comparisons were found by this review to be invalid, and a misrepresentation of the true results." (p2, paragraph 1).***

   This statement is not accurate as the "492%" is a different result for a different condition (helmeted baseline) than the "71%" (unhelmeted baseline) result. They are inherently not comparable.

   Moreover, the "492%"is an accurate representation of the true results, for the stated condition (i.e., the unhelmeted/CPD ATV in comparison to the helmeted/Baseline ATV). This comparison is valid and relevant, for the reasons stated in the immediately preceding comment.

In addition, as discussed in the immediately preceding comment, this statement could easily result in misunderstanding as both the single baseline results and the two baseline results were reported in Zellner, et al. (2012) (the former in the main text, the latter in Appendix I), as well as in the subsequent published papers (i.e., Zellner, et al. (2014a) and Zellner, et al. (2014b)).

The Zellner, et al. (2014a) paper discusses (in Appendix C) how both methods are valid, as the two baseline method gives the "intermediate results" (i.e., effects of a helmet; and effects of a Quadbar) while the single baseline method "can be considered to be a combination of the [intermediate] results" for the helmet and for the Quadbar.

The Zellner, et al. (2014b) paper discusses both the single baseline results and the two baseline results in the main text, with the overall conclusion being that "The CPD either does not have any statistically significant net benefits [i.e., the two baseline result], or else it is statistically significantly harmful [i.e., the single baseline result], depending on helmet use and on the baseline helmet/no helmet condition used for comparison."

This statement in Wordley (2012) is also inaccurate because the calculations, the methods and the comparisons in Zellner, et al. (2012) are accurate, valid and not "misrepresentations of the true results". As discussed in Zellner, et al. (2012) Appendix C, the 492% point estimate can be considered to be the result of the combined effects of not wearing a helmet (for which risk/benefit is large) and of fitting the Quadbar, which by itself was found to have a statistically insignificant effect.

It is reasonable and proper to consider the overall "combined" effect in comparison to the "intended use" (i.e., helmeted), which corresponds to the "single baseline" value, which in this example is 492%.

In addition, as discussed in Section V.43, this statement could easily give rise to misunderstanding as Wordley's criticism and this discussion is only relevant to the point estimate (or "sample mean") (i.e., the value preceding each set of square brackets in the Zellner et al. (2012) report). In contrast, the overall result reported by DRI in Zellner, et al. (2012), Zellner, et al. (2014a) and Zellner, et al. (2014b) is that, for the "population estimate" (i.e., the range denoted *within* each set of square brackets in the Zellner et al. (2012) report), the Quadbar CPD has *no statistically significant effect*, for either the helmeted (two baseline) or unhelmeted (two baseline) condition. In addition, as discussed in Zellner et al. (2014a), these two baseline results indicate a substantial probability that the device has a net harmful effect, in addition to the finding that the injury risks of the CPD are similar in magnitude to the injury benefits of the CPD, which is inappropriate for a safety device.

Therefore this statement in Wordley (2012) is inaccurate, incomplete and misleading.

The Wordley (2012) statements below:	(if corrected) should say:
"A risk/benefit ratio of 492% is quoted (much less safe) …"	"A risk/benefit ratio of 492% [255%,788%] (a population estimate that is statistically significantly more dangerous based on the results for the simulation sample) compared to a baseline with full-face helmet and no CPD …"
"… 71% (slightly safer) …"	"… 71% [41%, 135%] (a population estimate that is not statistically significantly safer or significantly more dangerous based on the results for the simulation sample) compared to baseline with no helmet and no CPD …"

10. "…according to DRI's updated research, the use of a CPD actually returns a risk/benefit ratio of 68% [42%, 114%] (slightly safer) for the unhelmeted condition, and 108% [69%, 169%] (marginally less safe) for the helmeted condition." (p2, paragraph 2)

This statement is inaccurate, incomplete and somewhat misleading because the terms "slightly safer" and "marginally less safe" (or, more accurately, "marginally more dangerous"), and the 68% and 108% values, *only* apply to the point estimates (e.g., the sample mean), which are based on the DRI *simulation sample* of 110 x 7 runs. This sentence should (if corrected) say "… a CPD actually returns an estimated risk/benefit ratio of 68% [42%, 114%] (a population estimate that is neither statistically significantly safer nor significantly more dangerous, based on the results for the *simulation sample)…*" and "… 108% [69%, 169%] (a population estimate that is neither statistically significantly safer nor significantly more dangerous, based on the results for the *simulation sample)*".

In summary, as clearly stated in the DRI report, the results for the Quadbar are statistically insignificant, *for both the helmeted and the unhelmeted condition*, i.e., using the "two baseline" method (i.e., helmeted/CPD ATV compared to helmeted/ baseline ATV; and unhelmeted/CPD ATV compared to unhelmeted/baseline ATV) that Wordley (2012) prefers.

Wordley (2012) cannot have it both ways: Wordley (2012) cannot simultaneously *reject* DRI's "simulation sample" as being invalid, which is stated and implied in the Wordley (2012) report,[11] and at the same time *accept the specific "sample mean"*

[11]Wordley (2012), page 1, paragraph 2 states that *"A number of serious limitations were identified* [in Wordley and Field (2012)] *with respect to the scenarios chosen and simulation methods utilised by the existing [DRI] research"*. The supposed limitations included those regarding the simulation sample used.

results that apply *only* to DRI's simulation sample, and not to the population estimate for all overturns.

As noted, the overall population results (as opposed to the point estimates based on a specific simulation sample) reported by DRI are that for the *population estimate* applicable to *all* samples of ATV overturns, the effects of the Quadbar are *statistically insignificant*, within a 95% confidence interval *for both the helmeted and the unhelmeted condition*. A 95% confidence interval (equivalent to p=0.05, Box, et al. (1978)) is by far the most widely accepted standard for being "somewhat convinced" of the reality of an outcome.

11. ***"Due to the fact that the 95% confidence intervals (shown previously in the square brackets) for both these results were observed to straddle the neutral risk/benefit value (100%), these findings were considered statistically insignificant." (p2, paragraph 2)***

    This statement is inaccurate as these results were not "considered" statistically insignificant. They *were and are* by definition statistically insignificant. This is a mathematical fact, given a 95% confidence interval (equivalent to p=0.05, Box, et al. (1978)) is by far the most widely accepted standard for being "somewhat convinced" of the reality of an outcome.

12. ***"Given no significant change in the updated DRI research and in consideration of other publications (initially reviewed) supporting the benefits of CPDs, the findings and recommendations of the original review have been reaffirmed." (p2, paragraph 4)***

    Both of the "givens" in this statement are inaccurate, inconsistent, incomplete and biased, as is the conclusion of this sentence. While it would be accurate to say that there was no change in the overall *results* found by DRI, between the DRI 2007 report and the DRI 2012 report (which both indicated that the Quadbar had no statistically significant effect for the population estimate), the "research" itself was certainly *substantially* changed, in that all of the previously raised supposed ***"serious limitations"*** mentioned by Wordley and Field (2012) were *fully* addressed, either by explaining them as being due to misunderstandings and/or false premises assumed by the critics or, if arguably valid, by fully accounting for them in the updated Zellner, et al. (2012) research. In short, there was a ***"significant change in the updated DRI research"***, in that all of the claimed ***"serious limitations"*** raised by Wordley and Field (2012) were fully addressed and found (by virtue of no change in overall results) to have been largely irrelevant to the overall outcome.

These earlier *"limitations"* had been fully addressed in Zellner et al (2012, 2012b), but no reference to this is made either in Wordley (2012) or Wordley and Field (2012). It is to be inferred from the failure to acknowledge Zellner et al (2012, 2012b) that Wordley continues to reject the validity of the simulation sample.

On the other hand, Wordley (2012) *does not address in any way* the major limitations of the second “given” in this statement, i.e., the Snook (2009) “ATV-alone” research, which is the only testing work (albeit faulty) allegedly ***“supporting the benefits of CPDs”***, a review of which shortcomings was directly accessible to Wordley on the internet as of 19 October, 2012, in the Zellner, et al. (2012b) report, several weeks before the Wordley (2012) report was finalized, and several months before it became available. In a rational process, being made aware of these major limitations should have changed the conclusions of Wordley (2012) regarding ***“consideration of”*** the Snook (2009) research, but unfortunately it did not.

The statement above indicates bias, in discounting the former substantially explained, refined, and updated DRI research, and completely ignoring the major limitations of the latter, severely flawed (“ATV-alone”) research, which Wordley (2012 states ***“support[s] the benefits of CPDs”***.

– ***“In the meantime, regulatory bodies should consider recommending the use of appropriately tested crush protection devices (CPDs) for riders who use Quad bikes at low speeds in the workplace and on farms.” (p5, 1st bullet).***

  This statement is inappropriate and based on a false premise, as discussed in the previous Section II.12, namely the *false premise* that the severely flawed Snook (2009) (“ATV-alone”) research “[validly supports] the benefits of CPDs”, and that the updated DRI research, which addressed all of the “significant limitations” previously raised in Wordley and Field (2012) , should for unstated reasons be totally ignored with regard to its overall findings that Quadbar CPD has statistically insignificant effects on the population estimates of net benefits in terms of injuries and fatalities (as well as a risk/benefit percentage that is larger than that for any known vehicle safety device).

– ***“Such devices have been shown to reduce the severity of Quad bike roll-overs, and have the potential to reduce the injuries and fatalities associated with these loss of control events.”(p3, paragraph 5)***

  This statement is unclear, without clear basis and could easily give rise to misunderstanding to mean that the Quadbar is actually effective in reducing injuries and fatalities. The statement that “such devices…reduce the severity of Quad bike roll-overs” is based entirely on the severely flawed (“ATV-alone”) work of Snook (2009), which mistakenly assumed that a more severe ATV rollover was one in which the Baseline ATV rolls *further away from* an already separated rider. Actually, rolling further away from an already separated rider is *less* severe, (not more severe), as it *decreases* the likelihood of rider/ATV contact. Likewise, the statement that “such devices…have the potential to reduce the injuries and fatalities” is based entirely on the same severely flawed Snook (2009) work, which mistakenly and without basis made assumptions regarding where an imaginary rider “would be”, for the baseline ATV and for the Quadbar ATV (with inexplicably different assumptions for each vehicle). This always assumed that the imaginary rider “would be” in a position in which the Baseline ATV always landed on him with an assumed injurious/fatal outcome, while with the Quadbar ATV the imaginary rider by some unknown mechanism always inexplicably remained in the “protective space” provided by the Quadbar (even though the imaginary rider was not restrained to the vehicle in any way). Similarly, although the Quadbar ATV had an increased tendency to come to rest on its side or on its wheels, the imaginary rider never ended up underneath it in these attitudes, even though there is ample real world evidence that fatal/serious injury entrapments/asphyxiation occur in these on-its-side and on-its-wheels ATV attitudes.[12]

[12] See Footnote 1.

SECTION III

COMMENTS ON "INTRODUCTION" (SECTION 1)

The comments on the Wordley (2012) "Introduction" (Section 1) are as follows:

1. ***"The 2011 [Wordley and Field (2012)] review found that:…Such [CPD] devices are generally designed to prevent the full weight of the Quad bike from being applied to, or coming to rest on the rider in the event of a roll-over." (p4, paragraph 2)***

   This statement that CPDs are "designed to prevent the full weight of the Quad bike from being applied to, or coming to rest on the rider" is without clear basis and based on a false premise that the rider will somehow stay on the ATV seat in an overturn. Common experience, plus the n=129 real world ATV overturn videos analyzed by Van Ee, et al. (2012), plus the 12 overturn tests with a passive crash dummy analyzed by Van Auken, et al (1998), all confirm that real riders and rider surrogates very *rarely* stay on an ATV seat during an overturn, and generally separate from the seat as it begins overturning (as a result of natural or active forces). This means that a CPD in fact *can* contact the rider, and *can* (e.g., if the rider is on the ground) concentrate and transmit the "full weight" of the ATV to the rider.

   Likewise, some CPDs (e.g., the Quadbar) are designed to, and in fact do, increase the likelihood of the ATV coming to rest on its side or on its wheels, which increases the likelihood of entrapment/asphyxiation/serious injury under those more load-concentrated surfaces. There is ample evidence that fatal/serious injury entrapments/asphyxiation occur in these on-its-side and on-its-wheels ATV attitudes.[13]

   In contrast, it is highly unlikely with a baseline ATV (without CPD) that the "full weight of the Quad bike" is ever "applied to" or comes "to rest on the rider", as the ATV surfaces are greater in extent than the rider surfaces. For example, such a hypothetical situation would require the ATV to be perfectly "balanced" on top of the rider, with *all* of its weight transmitted through some small area of the ATV. There are few if any conceivable realistic situations in which such a balanced, load-concentrated condition could occur (e.g., the ATV perfectly balanced and standing on its nose on a rider). Generally, if a portion of the rider is "beneath" an ATV, a portion of the ATV weight is born by the ground and a portion is born by the rider. The fact that there were no serious injuries or fatalities in the n=129 real world ATV overturns analyzed by Van Ee, et al. (2012), and in 79% of those overturns there was no injury at all (other than minor abrasions and contusions), despite there being many contacts between the ATV and rider, tends to confirm that injuries in ATV rollovers are relatively rare. That is not to say they are never injurious or never fatal. However, the exact human, vehicle, environment and motion conditions associated with serious injury and fatality in ATV overturns are not well enough understood at the current

[13] See Footnote 1.

time. Primarily, this is because of the inadequate data typically collected following ATV-related serious injury and fatality incidents.

2. ***"Several critical limitations were identified with these numerical models;"(p4, 1st bullet)***

To the extent that this statement refers to DRI research, it is inaccurate. None of the criticisms made of the models were, even if valid, "critical" to the results and the conclusions reached.

Moreover the statement was out of date when it was made: it does not take into account and appears to ignore the fact that all of the so-called "serious limitations" were either fully explained (e.g., as being due to misunderstandings or false premises assumed by the critics), or if arguably valid, were fully accounted for, in the Zellner, et al. (2012) research. This was fully explained in Zellner, et al. (2012b), which was directly accessible to Wordley on the internet[14] as of 17 October, 2012 (which was advised to ISCRR on the same date), several weeks before the Wordley (2012) report was finalized, and several months before it became generally available.

3. ***"These computer simulations generally modelled roll-over incidents based on 110 brief and largely incomplete accident descriptions drawn from the US and the UK. A large number of assumptions and interpretations were required, many of which had the potential to greatly affect the predicted injury outcomes;"(p4, 2nd bullet)***

This statement is inaccurate, unclear and could easily give rise to misunderstanding. First, as clearly stated in Zellner, et al. (2012), the simulations did not "[model] roll-over incidents", but, rather, modeled examples of several *general types* of overturns, based on the limited information available from real accidents, plus a systematically applied set of plausible assumptions. In contrast to what Wordley (2012) claims, none of the simulation runs were intended or assumed to be an exact recreation of a real accidental "roll-over incident", which is generally beyond the state of available information and the current state of simulation or testing technology.

Second, the characterization of the accident descriptions as "brief and largely incomplete" is a negative comment that can be made about virtually all accident databases, worldwide. For example, DRI staff members have participated – at the request of government agencies and manufacturer associations -- in on-scene in-depth accident investigations of thousands of motorcycle accidents in seven countries, each involving collection of approximately 2,000 variables[15]; however even those data could be considered to be "brief and incomplete" in comparison to all the information

---

[14] At www.dri-atv-rops-research.com

[15] Applying OECD, Motorcycles - Common Methodology for in-depth accident investigations, Parts 1 - 13, Paris, 2002, revised 2007.

that would be needed for *exact* modeling of real “roll-over incidents”, which was *not* DRI’s purpose in the ATV research.

Third, the statement that “large number of assumptions and interpretations were required, many of which had the potential to greatly affect the predicted injury outcomes” is a negative comment that could be made about almost any simulation model that exists and (in addition) is only relevant if DRI’s goal was exact modeling of a real accident, which was not the purpose of the DRI’s studies. The comment might have been relevant had DRI’s simulations not been correlated to very high level of agreement with physical tests of ATV/dummy overturns (with which DRI’s simulations have been correlated, in, for example, Van Auken (1998)). The same negative comments could be made about, for example, the simulations NASA ran for the Mercury, Gemini and Apollo programs, which, crude though they may have been in comparison to modern simulations, were close enough to enable successful near-earth space flight, lunar and earth orbits, inter-orbital mechanics, lunar landing, atmospheric re-entry and earth landing. Similarly, DRI developed and delivered the in-space “PILOT” simulator for use on-board the US Space Shuttle[16], which was mandatorily used on all Space Shuttle flights after 1994, to ensure Space Shuttle commanders were sufficiently “current” on skills for Shuttle final approach and landing to ensure safe landing after more than 10 days in space. “PILOT” was judged by NASA to be “close enough” to reality to be useful and mandatorily required (and was judged by NASA to be worthy of its Technical Achievement Award).

Instead, DRI’s purpose in the ATV research was to define a set of several *general types* of ATV overturns, based on the limited information available from real ATV accidents plus a systematically applied set of plausible assumptions. This was judged to be preferable to defining a set of purely hypothetical conditions bearing no relationship to the limited information that is known about real ATV overturn accidents.

All of these points are extensively discussed in Zellner, et al. (2012b) and Zellner, et al. (2012).

4. ***“[The 2011 review found that…] several inconsistencies and inaccuracies were identified in these [simulation] works and the interpretation and representation of the stated results.” (p4, 3rd bullet)***

This statement is unclear in regard to whether it is pertaining to, for example, differences *between* simulation researchers (e.g., Rechnitzer, et al. (2003), Grzebieta, et al. (2007), Zellner, et al. (2004)), or to differences *within* a given research report, or to differences within a *series* of reports. If it is referring to supposed “inconsistencies and inaccuracies” within DRI reports, then these have all been previously fully and extensively addressed in Zellner, et al. (2012), Zellner, et al. (2012a) and Zellner, et al. (2012b).

[16]DRI’s “PILOT” Space Shuttle in-space approach and landing simulator was awarded NASA’s Technical Achievement Award in 1994.

5. ***"A systematic review of these results indicated that the most promising CPDs had the potential to slightly increase the safety of Quad bike riders when subjected to roll-over for both the helmeted (recommended use) and non-helmeted (foreseeable use) conditions." (p4, 3rd bullet)***

   First, this statement is unclear with regard to which "results" it is referring.

   Second, since the only CPD simulation research that considered helmeted and unhelmeted conditions prior to this Wordley and Field (2012) review being discussed was the DRI research (e.g., Munoz, et al (2007) and Zellner, et al. (2008)), use of the phrase "had the potential to slightly increase safety" is inaccurate and is a mischaracterization of those results, which were in fact that the two subject CPDs (U-Bar and Quadbar) had *statistically insignificant effects* on injuries. In both cases, the confidence intervals were sufficiently wide that there were very large probabilities (i.e., for the helmeted case, p/2=0.35 and p/2=0.49, respectively, see Section V, Table 1.) that those devices would have *net harmful* effects in terms of population estimates (i.e., relatively large probabilities that they in fact had a *negative* net injury benefit for the population of all overturns).The Quadbar-helmeted condition, for example, had a risk/benefit of 99% [53%, 192%], a net benefit of 0%±21%, and a probability value (p-value) of 0.98. The latter can be interpreted as meaning that, if the Quadbar does have some non-zero effect, then the probability that the true net benefit is *negative* (i.e., that it is harmful in terms of increasing the normalized injury cost) given the simulation results, is 0.49 [=p/2]. This is nearly equivalent to a random "50/50" coin toss: heads (the device is harmful) or tails (the device is beneficial). It is *not reasonable, prudent or ethical* to recommend a device with this much potential for worsening injury outcomes.

   In addition, the findings of those two DRI studies were that both devices had risk/benefit percentages that were much larger than that for any other known vehicle safety device, and also larger than the ISO 13232 (2005) risk/benefit guideline. For Wordley (2012) to state that they "had the potential to slightly increase safety" is a gross mischaracterization of those results.

6. ***"Opinions were mixed as to whether the magnitude of this benefit could be deemed to be statistically significant;" (p4, 3rd bullet)***

   First, this statement is unclear with regard to whose "opinions were mixed". There are no clear references in the Wordley and Field (2012) review that indicate "mixed" opinions, or that there were any specific persons expressing any "opinions" in regard to the "statistical significance" of the results.

   Second, and importantly, "[f]indings are said to be *statistically significant* when the null hypothesis has been rejected. Thus, if results achieve statistical significance, the researcher concludes that treatment effect occurred." (Gravetter and Wallnau (1996), p 240). According to Parker (ed., 1984), the significance level is defined as "[t]he probability of false rejection of the null hypothesis." According to Box et al. (1978), p 109:

> "A series of conventional "critical" significance levels is in common use. These levels correspond to the probabilities representing varying degrees of skepticism. When the probability that a discrepancy as large as that observed, or larger, might occur is smaller than one of these critical probabilities, the discrepancy between the observation and hypothesis is said to be "significant" at that level. As a guide, it could be said that, when one's attitude is a priori "neutral" to a particular type of discrepancy, one begins to be slightly suspicious of a discrepancy at the 0.20 level, somewhat convinced of its reality at the 0.05 level, and fairly confident of it at the 0.01 level."

Of these levels, the p=0.05 level of significance (corresponding to a 95% confidence interval) is the most widely accepted upper bound on the level of significance applied in the statistical analysis field for decisions regarding whether an apparent non-neutral effect is real. For example, for p=0.05, if the true difference was zero, a deviation in either direction as large as that experienced, or larger, would occur by chance only about 5 times in 100, which implies that one can be "somewhat convinced" that the difference (e.g., due to a treatment, such as a CPD) is real.

In contrast, as noted by Box, et al (1978), at a significance level of 0.20, "one begins to be slightly suspicious of a discrepancy" being real, but that is all. There is still a substantial chance (i.e., 10%) that another sample would produce a discrepancy of at least the same magnitude as the observed discrepancy, *but in the opposite direction* (in this case, harmful rather than beneficial, or vice versa), when under the null hypothesis it is assumed that there is no real net benefit. In the presence of such a large uncertainty as to the direction of the outcome (i.e., harmful versus beneficial), it is traditionally considered not advisable to proceed with implementing such a treatment.

Since the U-bar and Quadbar sample means in these studies were both found to have p-values *substantially greater than 0.05*, then according to Box, et al. (1978), one cannot be "somewhat convinced" that that any net benefit suggested by that sample is real. This is not a matter for opinion; rather it is a matter of universal practice in the research field. Treatments having net benefit p-values substantially exceeding 0.05 (whether they be a medical device, pharmaceutical, medical procedure or safety intervention) are generally not released for clinical trials on humans, because one cannot be "somewhat convinced" that any net benefit suggested by that sample is real, or in fact is not harmful (i.e., have a negative net benefit).

7. ***"Experimental tests of a Quad bike fitted with a CPD (but without a crash dummy) and subjected to a range of different roll-over modes, velocities and terrain types found that CPDs could be effective in reducing the likelihood of complete roll-over of the bike." (p4, 4th bullet)***

First, this statement is not relevant, as the "likelihood of complete roll-over of the bike" has nothing to do with the likelihood of injury. There is ample evidence[17] that fatal/serious injury entrapments/asphyxiation can occur in ATV on-its-side and on-its-wheels final attitudes, as well as on-its-top. Conversely, "complete roll-overs" were observed in many or most of the n=129 overturns analyzed by Van Ee et al. (2012), however 79% of those overturns were reported as being non-injurious. Wordley (2012) (as well as the designers of the Quadbar) has mistakenly assumed, without basis, that preventing a "complete" (or "on-its-top") ATV rollover will somehow reduce or eliminate injuries and fatalities, whereas they are more likely to be simply transferred to other final attitudes (which is what was found in, for example, in Zellner et al. (2012)).

Second, the reference to a "range of different roll-over modes, velocities and terrain types" could easily give rise to misunderstanding. These "roll-over modes, velocities and terrain types" refer to the severely flawed tests of Snook (2009) which, among other things, involved sliding an ATV *sideways* down a ramp, and then tripping it into a rollover, both of these motions being highly unrealistic motions for off-road ATVs. This sideways slide-then-trip methodology has little to or nothing to do with ATVs, which if they overturn, typically do so while they are traveling forward or rearward on their wheels. The Snook (2009) "sideways" overturn tests were run with zero forward velocity and only relatively low (less than 11 km/h) lateral velocities. The "back flip" tests were run with less than 5 km/h of rearward velocity.

A further extensive discussion of the severely flawed Snook (2009) tests is provided in the Zellner, et al. (2012b) report.

8. ***"With a CPD fitted, there were no scenarios tested where the quad bike came to rest in a position which was considered to be more detrimental to rider safety than the bike without such protection." (p4, 4th bullet)***

This statement, taken from the Snook (2009) report, is misleading and based on several false premises. The fact that the use of the phrase "rested in a position that was *considered* more detrimental" (italics added) indicates the highly subjective and speculative method of evaluation used by Snook (2009).

As discussed above, Snook (2009) mistakenly assumed that an improved ATV overturn was one which resulted in a quarter-roll, and immediate arrest of the ATV; however, a quarter roll onto an ATV's side can be highly injurious[18]; and immediate arrest (i.e., nearer the rider) tends to *increase* rather than decrease the chances of ATV/rider contact (as it keeps the ATV in closer proximity to the rider, who tends to separate from the vehicle in an overturn). Snook (2009) likewise mistakenly assumed that a more severe ATV rollover was one in which the Baseline ATV rolls *further away from* the already separated rider, when in fact, an ATV rolling further away

[17] See Footnote 1.
[18] See Footnote 1.

from an already separated rider is *less* severe, as it decreases the likelihood of rider/ATV contact.

Similarly, Snook (2009) mistakenly and without clear basis made assumptions regarding where an imaginary rider "would be", for the baseline ATV and for the Quadbar ATV (with inexplicably different assumptions for each vehicle). This always assumed that the imaginary rider "would be" in a position in which the Baseline ATV always landed on him with an assumed fatal/serious injury outcome[19], while, with the Quadbar ATV, the imaginary rider always inexplicably remained in the "protective space" provided by the Quadbar, even though real ATV riders are not restrained in any way and tend to separate rapidly from the ATV in an overturn (either as a result of natural forces or active dismount), as can be observed in the Van Ee et al. (2012) videos.

Similarly, although the Quadbar ATV had an increased tendency to come to rest on its side or on its wheels, Snook (2009) mistakenly assumed that an imaginary rider would never end up underneath the ATV in these attitudes, even though there is ample evidence that fatal/serious injury entrapments/asphyxiations occur in these ATV on-its-side and on-its-wheels attitudes.[20]

9. ***"Based on this research, it appears likely that appropriately designed CPDs could lead to reduced injuries and fatalities for the common roll scenarios tested;" (p4, 4th bullet)***

   This statement is inaccurate as it is based on the false premise that the Snook (2009) tests were valid, when in fact (as discussed in Section III.7 and III.8 immediately above) those tests were severely flawed, highly speculative and based entirely on a series of mistaken assumptions.

10. ***"Quite significantly, this report returns vastly different predictions for the potential safety benefits due to the addition of the CPD compared to the earlier work. DRI's implied risk/benefit ratio for the use of a CPD by an unhelmeted quad bike rider increased from 71% in 2007, to 492% in the recent work."(p 5, 4th paragraph)***

    This statement is inaccurate and misleading. Wordley (2012) has compared results for "apples and oranges" to create a false impression of "vastly different" results. The 71% [41%, 135%] value is described in the Munoz, et al. (2007) report to be the unhelmeted *two baseline* result[21], while the 492% [255%, 788%] value is described in the Zellner, et al. (2012) report to be to the unhelmeted *single baseline* result[22], which are not directly comparable.

---

[19] Fatal/serious injury outcome is not what Van Ee, et al. (2012) found in n=129 real ATV (without CPD) overturn accident videos in which there were no fatalities/serious injuries and in which 79% of these overturns involved no injury.

[20] See Footnote 1.

[21] Table 2, Munoz, et al. (2007).

[22] Table 7, Zellner, et al. (2012).

The distinction between the single baseline results and the two baseline results is made clear in the Zellner, et al. (2012) report by stating what the baseline is in each case (e.g., "The risk/benefit percentage for the unhelmeted configuration is substantially higher (i.e., 492% [255%, 788%]) assuming the baseline is the baseline ATV with helmeted dummy)").[23] There were no statements in Zellner et al. (2012) comparing the two baseline result to (either the previous Munoz et al. (2007) or the updated Zellner et al. (2012)) single baseline results, as the single and two baseline results are fundamentally not comparable (i.e., they are apples and oranges). The presentation of the results clearly stated these single baseline conditions versus two baseline conditions.

In fact, comparing the "apples to apples" *two baseline* results, for the helmeted condition the Munoz, et al. (2007) results reported a risk/benefit percentage of 99% [53%, 192%] (which is a statistically insignificant outcome, i.e., according to Box, et al. (1978), one cannot be "somewhat convinced" that the device has a net benefit, or that the device in fact is not harmful for the population of overturns), while the Zellner et al. (2012) updated results were 108% [69%, 168%] (which is also statistically insignificant), and which is a relatively small difference from the former results. These are not "vastly different predictions".

For the helmeted condition, the Munoz, et al. (2007) results reported a risk/benefit percentage of 71% [41%, 135%] (which is statistically insignificant), while the Zellner et al. (2012) updated results were 68% [42%, 114%] (which is also statistically insignificant), which is a relatively small difference. Again, these are not "vastly different predictions".

In addition to clearly stating the baseline for each result, the updated results described in the main text of Zellner, et al. (2012) were stated to be for a *single baseline*;[24] and it was also stated[25] that the results for *two baselines* were given in Appendix I of that same report.

11. "Such analysis is particularly important given that both reports are categorised as "grey literature", in that they have been self-published by the authors' own company Dynamic Research, Inc. of Torrance, California." (p5, 4th paragraph)

This statement is misconceived and could easily give rise to misunderstanding. Wordley (2012) uses "grey literature" as if it was a negative term, but in actuality it is neutral or even a positive term. "Grey literature" is a term from within the field of library sciences that can be defined as:

> "The Fourth International Conference on Grey Literature (GL '99) in Washington, DC, in October 1999 defined grey literature as follows: "That which is produced

[23]Zellner, et al. (2012), page 66.
[24]Zellner, et al. (2012), page 64, first full paragraph.
[25]Op cit.

on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers."[26]

And:

"In the health sciences, grey literature is vital for developing a more complete view of research on a particular topic and for producing systematic reviews and other rigorous approaches to evidence synthesis. Grey literature can be a good source for data, statistics and for very recent research results. Because there's no publisher-enforced limitation on length, these reports can be much more detailed than the journal literature. And they can help to offset issues related to publication bias such as:

- Publication lag: Results of studies may appear in grey literature, such as conference proceedings, a year or more before they appear in peer-reviewed publications.

- Positive result bias: Study results that show a negative or no effect are published in scholarly journals less often than those that show a positive effect. Those negative results may be found by reviewing the grey literature.

Grey literature is particularly important in the area of health policy where health technology assessments, economic evaluations, health systems impact assessments and comparative effectiveness research are of special interest."[27]

And:

"Grey literature can be found in the form of reports, conference papers, posters or proceedings, policy documents, preprints, data sets, standards, translations, clinical trial data, factsheets, dissertations, committee reports, and more"[28]

And:

"**Grey literature** is informally published written material (such as reports) that may be difficult to trace via conventional channels such as published journals and monographs because it is not published commercially or is not widely accessible. It may nonetheless be an important source of information for researchers, because it tends to be original and recent. Examples of grey literature include patents, technical reports from government agencies or scientific research groups, working papers from research groups or committees, white papers, and preprints. The term "grey literature" is used in library and information science.[29]

[26]http://www.greylit.org/about (accessed 10 July 2013)
[27]http://guides.library.upenn.edu/healthgreylit (accessed 10 July 2013)
[28]Op. cit.
[29]http://en.wikipedia.org/wiki/Grey_literature (accessed 10 July 2013)

This Wordley statement implies that the reports are somehow deficient (because they are “grey”). However, as indicated in the above definition, “grey literature” includes “technical reports from government agencies or scientific research groups, working papers from research groups or committees”, all of which are widely relied upon sources. In addition, this definition states that “grey literature” can be “an important source of information for researchers, because it tends to be original and recent”. And, “In the health sciences, grey literature is vital for developing a more complete view of research on a particular topic” that avoids factors such as “publication lag” and “positive result bias” (i.e., only publishing results where the outcomes are statistically significant). The latter facts are particularly relevant to the Zellner, et al. (2012) and related research, as the outcomes (i.e., net injury benefits) were indeed *not statistically significant*. This means that (according to Box, et al. (1978)) one cannot be “somewhat convinced” that the device has a non-zero net effect (i.e., net benefit or net harm) on the population in question. As stated in the University of Pennsylvania website, “Study results that show a negative or no effect are published in scholarly journals less often than those that show a positive effect. Those negative results may be found by reviewing the grey literature.”

This Wordley statement could also give the wholly incorrect impression that the reports are not “widely accessible”. However, both reports have been and continue to be available via the DRI Library at info@dynres.com; and as stated in reference lists, Zellner et al. (2012) is publicly available on the internet at www.dri-atv-rops.research.com. To date, it has been downloaded by more than 800 readers worldwide, including by more than 170 readers in Australia and more than 340 readers located in the USA.

In any event, this statement is also now out of date, as two conference papers have since been published internationally, which describe and extend the work in Zellner et al. (2012), namely Zellner, et al. (2014a) and Zellner, et al. (2014b), the first of these being a fully peer reviewed journal article.

12. ***“Neither the initial report, nor the updated report have been subjected to a scientific peer review process.” (p5, paragraph 4)***

This statement is now out-of-date as, in fact, a condensed, detailed version of the updated report was submitted as a paper to an international peer reviewed conference, successfully completed the peer review process without substantive changes, was accepted for formal publication by two international engineering societies (i.e., the Society of Automotive Engineers International (SAE) and the Society of Automotive Engineers Japan (JSAE)), and was further nominated, reviewed and accepted as a journal article in both societies’ transactions, as Zellner, et al. (2014a).

As such it was the first fully peer-reviewed scientific paper in the field of ATV CPD injury risk and benefit analysis.

SECTION IV

COMMENTS ON "DEFINITIONS AND NOMENCLATURE" (SECTION 3)

The comments on the Wordley (2012) "Definitions and Nomenclature" are as follows:

1. ***"This review will examine the use of four wheeled, motorised bikes, having a straddle seat and handlebars. Such bikes are commonly referred to as either Quad bikes, or All-Terrain Vehicles (ATVs). For clarity and simplicity, this review will hence forth [sic] refer to these vehicles exclusively as Quad bikes. In instances where a vehicle has been described as an ATV by the original authors, the term Quad bike will be used in its place." (p7, paragraph 2)***

   The term "All Terrain Vehicle" (ATV) is a legally clear and unambiguous term in the US, where more than 90% of such vehicles are manufactured and sold, and as currently defined under ANSI/SVIA 1 (2010). The latter Standard involves approximately 28 safety-related design and performance requirements. In 2009, the US Consumer Product Safety Commission adopted ANSI/SVIA 1 (2010), including the associated definition of ATV, as a mandatory safety standard under the Consumer Product Safety Improvement Act (CPSIA).

   The term "Quad Bike", as defined here in an *ad hoc* way by Wordley (2012), is an informal slang term, which is far broader than "All Terrain Vehicle", and includes vehicles which do not necessarily meet these 28 safety requirements. Such "four wheeled, motorized bikes,[30] having a straddle seat and handlebars" may look like ATVs, but they do not necessarily meet the 28 safety requirements that ANSI/SVIA 1-compliant ATVs must meet under US law. Such non-ANSI/SVIA 1- compliant vehicles currently are reported to comprise between 5% and 15% of the Australian market.

   Therefore, "Quadbike" and "ATV" are not equivalent terms.

2. ***"A Roll Over Protective Structure or System (ROPS), is a [sic] external frame or structure which forms a compartment to protect the rider from injuries caused by vehicle overturns and to a lesser extent, collisions." (p7, paragraph 4)***

   This statement is inaccurate and could easily give rise to misunderstanding, as "collisions" are not part of the generally accepted functional definition of a ROPS.[31]

[30]Note: a "bike" is defined by the Oxford Dictionary as "a bicycle or motorcycle", where each of these in turn is defined as "a two-wheeled vehicle".

[31]e.g., SAE J2194, Roll-Over Protective Structures (ROPS) for Wheeled Agricultural Tractors, "3.2 Rollover Protective Structure (ROPS)—A cab or frame for the protection of operators of agricultural tractors to minimize the possibility of serious operator injury resulting from accidental upset. The ROPS is characterized by providing space for the clearance zone inside the envelope of the structure or within a space bounded by a series of straight lines from the outer edge of the structure to any part of the tractor that might come in contact with flat ground and is capable of supporting the tractor in that position if the tractor

ROPS are intended for occupant protection in overturns, not for occupant protection in “collisions”.

Certainly, off-highway vehicles such as side-by-sides (e.g., Recreational Off-Highway Vehicles[32]) that have ROPS are not intended to be operated on highways, and consequently are not equipped with frontal “collision” protection systems, highway lighting systems, directional indicators, etc. In particular, the belt systems and frontal structures for off-highway vehicles in general are not required to meet the requirements for those that are applicable to highway vehicles that have greater exposure to collisions.

Whether, for example, off-highway side-by-side vehicle ROPS and their associated belts can afford some protection in “collisions”, which may vary with specific vehicles and specific conditions, would be a matter of coincidence, and it is not a matter of definition, requirement or impact testing.

---

overturns. The mounting structure and fasteners forming the mounting connection with the tractor are part of the ROPS.” (Note: in Australia, ROPS are not required for tractors of less than 560 kg mass. ATVs are not tractors, and in addition, typically have a mass of around 275 kg.

[32] Anon., ANSI/ROHVA 1-2011., 2011.

SECTION V

COMMENTS ON "REVIEW OF DRI-TR-12-06" (SECTION 4)

The comments on the Wordley (2012) "Review of DRI-TR-12-06" (Section 4) are as follows:

"4.1 Overview of the initial research"

1. ***"A sample of 113 overturn events were initially simulated, these being loosely based upon brief and incomplete Quad bike overturn incident reports drawn from the UK's Health and Safety Executive (HSE) and the USA's Consumer Product Safety Commission (CPSC)." (p 9, paragraph 2)***

   This statement could easily give rise to misunderstanding. The characterization of the accident descriptions as "brief and incomplete" is a negative comment that can be made about virtually all accident databases worldwide. As only one example, DRI staff members have participated – at the request of government agencies and manufacturer associations -- in on-scene in-depth accident investigations of thousands of motorcycle accidents in seven countries, each involving collection of approximately 2,000 variables; however even those data could be considered to be "brief and incomplete" in comparison to all the information that would be needed for an exact recreation of a real accident, *which was not DRI's purpose*.

   Instead, DRI's purpose was to define a set of several *general types of overturns*[33] based on the limited information available from samples of real accidents, plus a systematically applied set of plausible assumptions for the unknown variables, based on testing and operational experience with this type of vehicle. This approach to develop a dataset of typical ATV overturn accidents was judged to be preferable to defining a set of purely hypothetical conditions bearing no relationship to the limited information that is known about real ATV overturn accidents.

2. ***"The authors acknowledged that, on average, only 8 out of 17 initial conditions were reported for these 113 overturn events, and no estimate or confirmation of the accuracy of these reports was provided." (p 9, paragraph 2)***

   This statement is inaccurate and could easily give rise to misunderstanding. DRI never referred to them as "only 8 out of 17 initial conditions" as, first of all, having some information about 8 initial conditions for each accident on average (in addition to the fact that an overturn was involved) is a *relatively substantial amount of information* about a given "type" of overturn, and far more than no information or purely hypothetical conditions, without any reference to accident data.

[33] Also referred to as "generic roll scenarios" by Wordley (2012) in his Section 6 "Recommendations".

Second, 8 known initial conditions (as stated) was an average, and the US/CPSC cases on average had information on more than 8 initial conditions, whilst the UK/HSW cases on average had information on fewer than 8 initial conditions.

Third, the statement that "no estimate or confirmation of the accuracy of these reports was provided" is a negative comment that can be made about virtually all accident data databases worldwide. Yet, accident databases are very useful in defining typical conditions for, for example, crash testing, and for computer simulations involving typical crash conditions.

3. ***"The researchers devised their own estimates for all unreported initial conditions, with the only criteria [sic] being that an overturn event resulted. The size of this scenario data set was then also artificially increased via the generation of an additional six cases for each reported case, via unspecified perturbations made to the initial conditions of each case." (p 9, paragraph 2)***

This statement is inaccurate in several regards, including in regard to "the only criteria [sic] being that an overturn event resulted." In the earlier DRI research (to which this statement is referring), twenty eight (28) *other criteria* were systematically applied, as described in Appendix B of Zellner, et al. (2012). Beyond resulting in "an overturn event result", other criteria included compatibility with the known information, and plausibility based on DRI's more than 20 years of testing and operational experience with this category of vehicle in diverse environments. It is inaccurate to say that the fact that an "overturn event resulted" was the "only criteria" used.

The statement that in DRI's previous research, the scenario dataset was "also artificially increased…via unspecified perturbations" is inaccurate and can also be easily misinterpreted as being somehow inappropriate. First of all, the perturbations were not "unspecified" but are described in Table 4 of Zellner, et al. (2012).[34] Second, perturbation analysis (or sensitivity analysis) is a widely accepted and applied method that can be used to take into account variations in initial conditions (which variations can be considered to represent uncertainty, as is the case in all accident investigation information). This method also tends to reduce the sensitivity of the outcome to the specific conditions found in a relatively small number (i.e., n=113) of overturn types and increases the statistical degrees of freedom of the analysis. This can then increase the statistical significance of the outcomes (the limitation thereof which was one of the previous criticisms directed toward the Munoz, et al. (2007) analysis which used "only" n=113 types of overturn).

[34]This same Table 4 appeared in the draft of the DRI FISITA (2008) paper (incorrectly cited in Wordley (2012)) that was submitted to the FISITA conference. The Zellner, et al. (2008) FISITA paper was the first time the perturbation method was used, in order to attempt to improve the statistical significance of the results, and to reduce the sensitivity to the specific conditions found in a relatively small number (i.e., n=113) of overturn types. However, the FISITA paper reviewers requested that the paper be substantially shortened, due to length restrictions, so this Table 4 was removed from the final published proceedings paper. Table 4, which further describes the FISITA paper methods, does appear, as noted, in the Zellner et al. (2012) report.

An explicit listing of the perturbed initial conditions would have required 20 pages of tabular data, which was judged to be excessive. An explicit listing, for both the earlier analyses and for the "Updated analyses" can certainly be provided to qualified researchers on request.

Note that in the latest "Updated analyses" reported in Zellner, et al. (2012), the perturbations can be *exactly* reproduced by anyone by applying the algorithms in Appendix B of Zellner et al. (2012) to the n=110 sets of initial conditions defined in Appendix C of Zellner, et al. (2012). An explicit listing of perturbed initial conditions is therefore superfluous.

4. ***"These initial simulations predicted a 99% risk/benefit ratio (i.e. a marginal improvement in safety) for helmeted riders due to the addition of the Quad Bar, and a 71% risk benefit ratio (i.e. a more significant improvement in safety) for unhelmeted riders." (p 10, paragraph 2)***

This statement is incomplete, inaccurate and could easily give rise to misunderstanding. This statement is incomplete because the results described in Munoz, et al. (2007) were a risk/benefit percentage of 99%[53%, 192%] for the helmeted condition, and a risk/benefit percentage of 71%[41%, 135%] for the unhelmeted condition. The values in the square brackets indicate the 95% confidence interval for the population estimate and are as important (if not more important) to the meaning of the results as the point estimate (i.e., "sample mean") value in front of the square brackets. In both the helmeted and unhelmeted conditions, the "complete" results (including the confidence intervals) indicate that the risk/benefit outcome is *not statistically significantly different from 100%*, meaning that the injury risks of the Quadbar are effectively equal to the injury benefits; and that there is a substantial probability that the device has a harmful effect on the population of overturns.

A further analysis of these results is shown below in Table 1.

Table 1. Normalized Injury Cost risk/benefit, net benefit, confidence intervals and p-values for Quadbar ("V-Bar") and U-Bar ROPS, based on data from Munoz, et al. (2007) and Zellner, et al. (2008), respectively

Configuration	Risk/Benefit	(95% CI)	Net Benefit	(95% CI)	Probability value
V-Bar, helmeted	99%	[53%, 192%]	0%	$\pm$21%	0.98
V-Bar, unhelmeted	71%	[41%, 135%]	8%	$\pm$15%	0.31
U-Bar, helmeted	95%	[74%, 122%]	1%	$\pm$7%	0.69
U-Bar, unhelmeted	91%	[73%, 113%]	3%	$\pm$6%	0.40

One of the key results in Table 1 is the "probability value" (or "p-value") in the last column, which is the probability of the observed results occurring if the device does not have any effect on the Normalized Injury Cost net benefit (i.e., the Null

Hypothesis). If the p-value is less than 0.05, then the Null Hypothesis is considered to be unlikely and therefore the net benefit of the device is considered to be statistically significantly different than zero, in either the positive or negative direction. In contrast, if the p-value is larger than 0.05 then we cannot reject the Null Hypothesis and the result is considered to be not statistically significantly different from zero. However, if the Null Hypothesis that the net benefit is 0 is not true, and instead we assume that the true net benefit has some unknown value from a random distribution, then the probability that the observed and true net benefit have opposite signs is equal to the p-value divided by 2

.

Since the U-bar and Quadbar sample means in these studies were both found to have p-values substantially greater than 0.05 (i.e., ranging from 0.31 to 0.98), then according to Box, et al. (1978), one certainly *cannot* be "somewhat convinced" that that any non-zero net effect benefit (i.e., either net benefit or net harm) suggested by that sample is real. This is not a matter of opinion; rather it is a matter of standard practice in the research field. Treatments having net benefit p-values substantially exceeding 0.05 (whether they be a pharmaceutical, a medical device, a medical procedure or a safety intervention) are most often not released for clinical trials on humans, because one cannot be "somewhat convinced" that any net benefit suggested by that sample is real, or in fact that the treatment is not harmful (i.e., having a negative net benefit).

Furthermore, these results can also be interpreted as indicating that if the net effect is not zero, then the p-values ranging from 0.31 to 0.98 can be interpreted as indicating that the probability that the devices are actually harmful (=p/2) ranges from 0.155 to 0.490. These are large probabilities that the device is actually harmful. The 0.155 probability is nearly equivalent to chance that a roll of one dice will be "1" (the device is harmful overall) vs 2, 3, 4, 5, or 6 (the device is beneficial overall). The 0.490 probability is nearly equivalent to a random "50/50" coin toss: heads (the device is harmful overall) or tails (the device is beneficial overall). It is not reasonable, prudent or ethical to recommend a device with this much potential to worsen overall injury outcomes.

This statement in Wordley (2012) is inaccurate and could easily give rise to misunderstanding because, as noted above, the more complete results for the Quadbar of 99% [53%, 192%] (p=0.98) (helmeted) and 71% [41%, 135%] (p=0.31) (unhelmeted) from the Munoz, et al. (2007) study mean that the risk/benefit outcome for the population estimate is *not statistically significantly different from 100%; and that there is a substantial probability (i.e., 49% for helmeted) and 16% for unhelmeted) that, for the population of overturns, the Quadbar is in fact harmful*. The results categorically *do not mean* that there is a "marginal improvement" or "a more significant improvement in safety" as mistakenly stated by Wordley (2012). In fact, the 99% and 71% values cited by Wordley are *only* "point estimates" that apply to the sample means. The point estimates are *only* relevant if one assumes that each and every ATV overturn in the population of all ATV overturns can be exactly modelled by one of the 113 types of overturns that were simulated, and that each of the 113 types of overturns represents N overturns in the population (e.g., N=1 would correspond to a census of all overturns), which clearly Wordley (2012) does not assume.

For those who consider the n=110 UK/US sample of overturn types to be a "complete census" of all overturns (which Wordley (2012) does not), and that therefore the results can be viewed deterministically (rather than statistically), and ignoring all other similar samples of overturns, then 99% and 71% risk/benefit percentage may represent a marginal improvement for this particular specific sample of overturns, but still having risk/benefit percentages which are:

> "…much higher than those found in any published data for any automotive safety device, for which the injury risks have been found to be less than 7% of the injury benefits; unacceptably high injury risks in comparison to injury benefits in comparison to the suggested reference guideline for research purposes published in ISO 13232 (2005), i.e., the Quadbar injury risk/benefit percentages were all far greater than the *"not… more than 12 percent"* reference guideline indicated in International Standard ISO 13232-5 (2005); and unacceptably high injury risks in comparison to injury benefits, when compared to the "regulatory policies of several of the [Australian] states," mentioned in the Heads of Workplace Health and Safety Authorities (HWSA) Technical Engineering Group (88
> report, which were apparently stated to be that the "the benefits need to be at least 2 times the risks" (i.e., 2 deaths prevented for every 1 new death caused by a device, or 50 percent risk/benefit".[35]

5. ***"Overall these findings were dismissed as statistically insignificant, on the basis that the calculated 95% confidence intervals for both results straddled the neutral risk/benefit baseline." (p 9, paragraph 2)***

This statement is inaccurate and could easily give rise to misunderstanding, as the results were not "dismissed" as if there was some sort of subjective judgment being applied. The results were in fact *not statistically significantly different from 100%.* This is based on a 95% confidence interval, which is by far the most widely accepted and applied confidence interval used in scientific and statistical analyses.

As discussed in Section V.4 and Table 1 above, these 95% confidence intervals which straddle zero are associated with unacceptably large probabilities that the net benefit of the Quadbar is negative (i.e., harmful) (i.e., the data in Table 1 indicate that Quadbar has a 49% probability of being harmful for the helmeted condition, and a 16% probability of being harmful for the unhelmeted condition).

[35]Zellner, et al. (2012), page 76.

"4.2 Modifications and additions made in the updated work"

6. ***"The most significant changes include: …a tightening of the requirements for the AIS coding of reported injuries, resulting in a filtering of the reported injury data set;" (p 10, 4th bullet)***

   This statement could easily give rise to misunderstanding by the implied suggestion that an inappropriate procedure was somehow involved. The "significant changes" to the injury codings were required and appropriate because the specific purpose of the subject research was *to enable calibration of the simulation model*, which at the current state of technology is only able to predict a limited set of locations, types and severities of injury (as is the case for all crash dummies applicable to cars, motorcycles, etc., worldwide). It would therefore be *inappropriate to attempt to include those injuries that cannot be monitored by a crash dummy*. To include those injuries that cannot be monitored would be to attempt to compare "apples to oranges". This is the reason that some locations, types and severities of injuries were, appropriately, "filtered" (i.e., removed) from consideration.[36]

7. ***"The most significant changes include…a change to the baseline(s) used to normalise and present the revised injury risk/benefit results;"(p10, 5th bullet)***

   This statement is incomplete and could easily give rise to misunderstanding as involving a change to only one alternative "single baseline" method. In fact the previously used two baseline [i.e., separate helmeted and unhelmeted]) method "used to normalize and present the revised injury risk/benefit results" were *also* presented in Zellner, et al. (2012), Appendix I. The main text discussion (i.e., pages 65 to 68 of Zellner, et al. (2012)) is stated to be based upon the "baseline ATV with helmet" using the "single baseline" method, and that main text discussion also cross-references (i.e., on page 64) the presentation of results using the previous "two baseline method (i.e., "Supplemental, 2-baseline comparison tabulations are *presented* in Appendix I, for reference" (italics added)).[37]

   Moreover, this comment is also now out-of-date, as the more recently published papers describing the Zellner, et al. (2012) research present both the single baseline results and the two baseline results. The Zellner et al. (2014a) paper (in which single baseline results are presented in Appendix B and two baseline results are presented in Appendix C) and the Zellner, et al. (2014b) paper (in which both single baseline and two baseline results are presented in the main text) present results in both formats.

   As discussed in Section IV, Comment 37 below, the single baseline is relevant and important way to report the results (though not the only way the results were reported). This is because:

[36] In contrast, a general study of what types of injuries ATV riders may sustain (including non-dummy-monitor-able injuries such as abrasions, contusions and lacerations; as originally coded in DRI-TM-04-77 (December 2004)) might be of interest for some other research purposes, but this was not the purpose of the subject research, which was to calibrate the computer simulation.

[37] Zellner, et al. (2012), page 64.

– The "change to the baselines" is not a "significant change", as the updated results presented in Appendix I of Zellner, et al. (2012), which is based on DRI's previous reporting two baseline method, indicate the Quadbar has no statistically significant net benefit, and its risk/benefit is not statically significantly different from 100%, which is much greater than the risk/benefit for any known vehicle safety device, which is *very similar to the overall result* reported in Munoz, et al. (2008). Wordley (2012) elsewhere confirms there is ***"very little change"*** in the overall results using the "two baseline" method,[38] but the above statement describes the new "single baseline" reporting method as a "significant change";

– the single baseline method, though not the only method reported, expresses what may occur if riders who currently wear helmets mistakenly remove them under the mistaken assumption that with a CPD, helmets are not needed. This is an important scenario (among several) that may occur and is therefore relevant;

– the "single baseline" results are not "misrepresented", as the results reported in the main text of Zellner, et al. (2012) clearly state that they are for the "baseline ATV with helmet"; and as noted above, the main text cross references the results using the previous "two baseline" method.[39]

– the "single baseline" method is the more appropriate method to include in the main text because the "two-baseline" approach can be misleading in that it shifts focus away from a primary hazard, which is (in this case) not wearing a helmet. Not wearing a helmet is approximately 2.5 times more hazardous than wearing a helmet, and even if a device such as a CPD reduced that hazard by a small amount – which the DRI results for the Quadbar do not indicate – the unhelmeted hazard – and conversely the large benefits of wearing a helmet – *far exceed* the estimated effects of fitting a CPD (while unhelmeted), *which were found to be statistically insignificant*;

– the "single baseline" method is the more appropriate method to include in the main text because the "single baseline" results (presented in the main text of Zellner, et al. (2012)) can be considered to be a *combination* of the (previously used) two separate "two baseline" results (presented in Appendix I of Zellner, et al. (2012)). Therefore the "combined" result is appropriately reported in the main text, and the two "component" results are appropriately reported in Appendix I. This combined effect is further described in Zellner et al. (2014a).

[38] Wordley (2012), page 20, "The modified test methodologies and models used in the updated 2012 research have resulted in very little change to the overall risk/benefit results, based on the more appropriate multiple baseline comparison [as reported in Zellner, et al. (2012) Appendix I]."
[39] Zellner, et al. (2012), page 64.

– other reasons why the "single baseline" method is the more appropriate method to use in the main text are enumerated in the main text[40] and in Appendix B of Zellner, et al. (2012).

– as noted therein, the "two baseline" method, which was previously the only method used, can be potentially misleading (and consequently harmful) to the public. For example, US first generation airbags were originally only regulated, tested and tuned so as to be a "net benefit" (and a low risk/benefit) for one "misuse" condition (i.e., unbelted mid-size adult males) whereas it was later discovered that such over-tuning for this baseline resulted in a high risk/benefit percentage for another (i.e., alternate baseline) condition (i.e., small front seat passengers), resulting in approximately 380 airbag-induced fatalities to the latter category as of 1998.[41] As stated elsewhere, using a "single baseline" references the results to a single (ideally the intended) use, and avoids misinterpretation of such confusing and contradictory "multiple baseline" outcomes. This is also relevant to the ATV CPD case, as discussed immediately above, as the effect of wearing a helmet (i.e., first baseline) is much larger than and overwhelms any effect of the CPD (second baseline), an outcome which is not evident from the "two baseline" results.

– this statement by Wordley (2012) is now out of date, as Zellner, et al. (2014a) and Zellner, et al. (2014b) describe both the single baseline outcomes and the two baseline outcomes in the main conclusions.

8. ***"Clarifications are provided with respect to: …the methods used to generate guesses for incident initial conditions which were not provided by the incident descriptions;" (p 10, 7^th^ bullet)***

First of all, this statement is inaccurate in incorrectly describing the n=110 general types of overturn as "incidents". As discussed numerous times throughout Zellner, et al. (2012), the n=110 general "types" of overturn are just that, and they are not intended to represent specific "incidents" for which there was far too little information available in the case descriptions.

Each of the n=110 general "types" of overturn was defined based on, first of all, the limited available information for a given specific case (i.e., which provided an average of 8 out of the 17 variables needed to define an overturn), and beyond that, plausible, uniformly applied, systematic assumptions (rather than "guesses") that describe the other initial conditions of each general type (rather than specific "incident") of overturn. The refined plausible assumptions, defined in Appendix C of Zellner, et al. (2012), in turn were based on engineering experience gained through testing and operation of ATVs and awareness of others' research in this area during the last 30 years, e.g., indicating that typical (but not all) ATV overturns occur at relatively low speeds at or below 30 km/h (18 mi/h) and can occur at walking speeds

[40]Op. cit.
[41]Iijima, 1998.

(e.g., 5 km/h, (3 mi/h)) on steep slopes, etc.; and that "long slope" multiple overturns could plausibly be less frequent than had been originally assumed.

4.2.1 Changes to the dummy, bike and ground interactions and simulation domain

9. "The mechanics and interactions of the proprietary ATB dummy simulation were modified in the updated work, apparently in an attempt to address some of the limitations identified by critics of the original research." (p 11, paragraph 2)

This statement could easily give rise to misunderstanding by implying that use of a "proprietary" simulation code is somehow an issue. First of all, virtually all modern simulations codes in wide use are "proprietary" (e.g., MADYMO, LS-Dyna3D, etc.) so that is obviously not a negative factor. Secondly, DRI's "proprietary" version of the US/DOT-US/Air Force "ATB" simulation code was developed out of necessity for purposes of collision modeling of complex crash dummy and whole vehicle motions, owing to the large number of interacting body parts and contact surfaces in such scenarios. Moreover, the relevant enhancements were implemented for DRI by several dynamics experts including Dr. John Fleck, who was an original developer of the US/DOT-US/Air Force version of "ATB". The primary enhancements involved adding a larger number of additional masses, and fracture-able joints. Third, DRI has licensed its proprietary version of ATB to others worldwide, including manufacturers and government agencies, which have relied upon it; and the source code can be examined on request by qualified researchers in the presence of DRI representatives.

This statement could also easily give rise to the incorrect understanding that only "*some* of the limitations identified by critics" had been addressed, when in fact, *all* of the alleged limitations were either fully explained (e.g., as being due to misunderstandings and false premises assumed by the critics (e.g., as explained in, e.g., Zellner, et al. (2012a) and Zellner 2012b)), or if arguably valid, were fully accounted for in the refined simulations which were used in the Zellner, et al. (2012) research.

10. "Updated frictional coefficients for the dummy/soil interactions were also measured and incorporated into the new research, with the aim of addressing prior criticisms relating to the potentially unrealistic way in which the bike and rider bounced in previous simulations." (p 11, paragraph 5)

This statement is inaccurate, as the purpose of updating the dummy/soil interactions was not to address the alleged "unrealistic way in which the bike and rider bounced in previous simulations" but rather to address what was eventually discovered to be excessive forces and moments being applied to the head, neck and lower extremities by excessive dummy/soil friction, resulting in increased frequency of predicted injuries, as described in Zellner et al. (2012), pages 2 and 58.

The alleged "unrealistic" bounce was partially a result of a lack of understanding by the critics (in this case Lambert and his "beach ball" comment, as discussed in Zellner, et al. (2012a), and the previous "high energy" assumption, which was resolved by means of the updated "low energy" assumption (i.e., page 48 and following, as well as Appendix C, of Zellner, et al. (2012)), by imposing an maximum slope length of 3 metres.

11. ***"This change [involving adding a terminal horizontal plane, which was imposed three metres vertically below the starting point of the Quad bike was reported to limit the extent and number of rolls that both the bike and dummy would be subjected to, as previous research frequently utilised potentially unrealistically long slopes (perhaps 50m)." (p 11, paragraph 6)***

This statement is inaccurate, in regard to "potentially unrealistically long slopes". The original simulation models were originally developed for 3 sec simulation run times (e.g., Van Auken et al. (1998)), for which it was assumed that the ground slope was locally linear (i.e., planar). This locally linear assumption may not be valid for the longer simulation run times (e.g., 10 sec) used in the current study, and therefore a horizontal ground plane was added to the simulation models to impose a maximum slope length of 3 metres.

Long slopes have also been observed in Australian ATV accidents[42]. In fact, the prevalence of very long slopes in NZ ATV overturns was reported as a primary factor for the T-Bar CPD design, i.e., in order to arrest the continuous rolling motions of ATVs on very long slopes, to aid in recovery (and reduced rollover damage to) the ATV[43] which is independent from (and which to some extent can exacerbate) injury prevention.[44]

12. ***"These videos were not made available for this review." (p 11, paragraph 7)***

This statement could easily give rise to a misunderstanding that DRI has somehow resisted examination of its research. In fact, unlike all other researchers in this field, in the previous DRI research, DRI made available (during the course of the 2002-2009 Victorian inquests) *all* of its archive of several *hundred* simulation input files, *all* of its output files including several hundred videos generated from animations software and recorded on DVD (i.e., those animations from Van Auken et al. (1998) and from Munoz et al. (2007)), and its proprietary simulation software.

---

[42]For example, the Shepherd and Crole fatalities in the 2002-2009 Victorian Coronial inquest involved very long (15 and 20 degree, respectively) slopes. Evidence was submitted to the inquests that in such circumstances an ATV that was fitted with a ROPS and belts that increased the likelihood of continuous rolling would increase the likelihood of very high brain accelerations and therefore increased likelihood of brain concussive injuries.

[43]Arresting continuous rolls down long slopes and reducing ATV damage from such occurrences can be a benefit of some types of CPD; however, arresting the rolling motion of the ATV also tends to place the ATV in closer proximity to the separated rider, which is a hazardous (not a safer) trend. These opposing effects of CPDs (i.e., reducing the damage to and difficulty of recovering ATVs, versus arresting the ATV in closer proximity to the rider (as found in, for example, Zellner, et al. (2012, page 70) have long been a source of confusion in the discussion of the effects of CPD-type devices on ATVs.

[44] For example, arresting the ATV rolling motion immediately (i.e., in close proximity to where a rider is laying on the ground) can increase the chances of entrapment and/or contact between the ATV and rider.

Following internet publication of the DRI updated (i.e., Zellner, et al. (2012)) research report, which included a much larger number (i.e., several thousand, i.e., 3,080) of overturn simulations, Wordley's ISCRR institute inquired as to the availability of simulation "videos". DRI sent the following reply in October 2012, which is still valid and applicable:

> *"At the current time, there are no "videos" of the [updated] simulations. Computer-generated visual displays (i.e., "animations") of the 3,080 simulations can be generated from the output files, and they have been reviewed in detail, but they require licensed software to view.*
>
> *Creating videos would involve a significant effort, and extensive (i.e., perhaps terabytes) of storage capacity.*
>
> *More importantly, currently, we are not aware of any suitable scientific forum or mechanism for scientific exchange in this area. If and when such a forum and mechanism exists, we are willing to provide access to the simulations, to appropriately qualified researchers."*

In short, making videos of the simulations "available" is not a small thing, and would require hundreds of DVDs and an extensive effort to generate them. DRI is willing to make the underlying "visual displays" available to qualified researchers for inspection, if/when there is a suitable scientific forum. Likewise, for such a forum, DRI is willing to provide example videos drawn from the 3,080 simulations.

4.2.2 The addition of a new asphyxiation injury criterion

13. *"Without reviewing the relevant simulation videos it is difficult to generalise further about the reasons for this result." (p12, paragraph 2)*

This statement is misconceived, as no simulation "videos" exist. See response above, for Section V.12.

In addition, understanding what is visible in the simulation "visual displays" is highly dependent on the viewing angle chosen, and an understanding of what information is in the associated (but not explicitly modeled) UK/US case. Moreover, as extensively discussed in Zellner, et al. (2012b), previous comments and efforts to "generalize further" based on viewing videos from the previous research were very often inaccurate, misinterpreted, without clear basis, irrelevant, or based on false premises. As noted in Section V.12 above, however, in principle DRI is willing to provide access to the animations and/or to generate example videos.

14. *"Their report of entrapments occurring underneath the side of the bike is a relatively new phenomena [sic] compared to the initial study, and may be an artefact of the passive rider model used, the new "low energy" scenarios tested*

(discussed in more detail in Chapter 4.2.5), or the changes made to the dummy hand grip force." (p12, paragraph 2)

That description of "entrapments occurring beneath the side of the [ATV]" appeared in the Zellner, et al. (2012) report as a result of the inclusion of a preliminary asphyxia/breathing difficulty criterion in the latest DRI research; and a result of DRI investigating further and attempting to better understand the reasons for the number of "breathing difficulty" simulation outcomes being unchanged as a result of fitting the Quadbar.

This phenomenon is generally consistent with the claims of Snook (2009) as well as Lambert, Robertson and others that the Quadbar reduces chances of an ATV coming to rest upside down; and consequently more often coming to rest on its side (or on its wheels), both of which being known from ATV accident files (e.g., in the US, UK, AUS and NZ) to be associated with asphyxiation/serious injury.

The statement that "entrapments" may be a result of "the passive rider model used" is speculative and without clear basis, as "entrapments" have been reported in Australian and NZ *real* fatal and serious injury cases in the past, where presumably the rider was, at least to some extent, "active".

It is further noted that, according to Van Ee (2012), in about 28% of real ATV overturns involving real riders observed in n=129 internet videos, the rider did not attempt active dismount, and therefore may have behaved effectively as a "passive" rider during and after the overturn event.

Similarly, the statement that "entrapments" may be a result of "the new low energy scenarios tested" or "changes made to the dummy hand grip force" are also speculative and without clear basis, as "entrapments" were never analyzed in the previous DRI research, so that no comparison is possible in regard to whether "entrapments occurring underneath the side of the bike" occurred in the previous research.

In order to be able to make such statements, one would have to conduct a detailed cause-effect simulation-based analysis; along with systematically examining the simulation outcomes (in terms of entrapped/asphyxiated) from the previous research, in comparison to those from the updated research.

15. ***"Video review is definitely required to determine if such entrapments appear realistic, or if a real person might have easily and instinctively avoided such injury cases." (p 12, paragraph 2)***

This statement is speculative and without clear basis as, certainly, an ATV coming to rest on-its-side (or on-its-wheels) on real riders has been observed to result in asphyxiation/serious injury to real riders.[45]

[45]See Foot note 1.

In addition, this statement is speculative as it is unclear on what basis it would be subjectively judged (using "video review") whether the entrapments "appear realistic, or if a real person might have easily and instinctively avoided such injury cases." If such subjective judgment is as inaccurate as were the previous subjective judgments made in regard to the videos from the earlier DRI research (they being very often inaccurate, misunderstood, without clear basis, irrelevant, or based on false premises, as extensively discussed in Zellner, et al. (2012a)), then it is unclear how such subjective judgment would contribute ***"to determine if such entrapments appear realistic, or if a real person might have easily and instinctively avoided such injury cases."***

Nevertheless, as stated in Section V.12 above, DRI is willing to make the underlying "visual displays" available to qualified researchers for inspection, or example videos available, if/when there is a suitable scientific forum.

"4.2.3 New face and skull fracture mechanisms for unhelmeted riders"

16. "The assumption of the use of a full face, on-road style motorbike helmet represents a serious disconnect with contemporary Australian Quad bike helmet wearing practices, and more significantly, the recommendations being made by Australian regulatory bodies and even Quad bike manufacturers." (p12, paragraph 5)

This statement is inaccurate in several ways and represents a major misunderstanding in Wordley (2012).

First, the full-face Bieffe B12 helmet is fully compliant with the impact, retention, penetration and coverage requirements of AS/NZS:1698.[46] Second, AS/NZS:1698-2006 does not prescribe any particular style/type of helmet. Both full-face and open-face motorcycle helmets that are universally used by motorcyclists throughout Australia comply with AS/NZS:1698, as do some half-helmets (also known as "shorty" helmets), such as the THH T70 suggested by Wordley (2012). Third, the ATV manufacturers' recommendations regarding helmets are that they comply with the applicable federal, state and/or other standards[47] and are consistent with the guidelines provided by the ATV Safety Institute as follows:

> "Select a helmet that meets or exceeds your state's safety standards and carries either the Department of Transportation (DOT) label or the Snell Memorial Foundation label. Your helmet should fit snugly and fasten securely. Full-face helmets help protect your face as well as your head. Open-face helmets are lighter

[46]See Footnote 2.
[47]See Footnote 3.

> and may be cooler, but should be used with mouth protection. Eye protection should be used with both types of helmets."[48]

Note that the only ATV helmets illustrated in the aforementioned ATV Safety Institute guideline are full-face helmets.

Fourth, inquiries indicate that, in fact, neither the Federal Chamber of Automotive Industries (FCAI) nor the largest Australian ATV importer specifically recommend the "shorty" helmets which are asserted by Wordley (2012) to be the subject of "recommendations made by…Quadbike manufacturers". This comment has no known basis.

17. ***"The helmet standard which currently applies to Australian Quad bike users is AS/NZS 1698:2006, and an example of the style of helmet which passes this standard is shown in the figure below." (p 12, paragraph 5)***

This statement could easily give rise to misunderstanding since it wrongly implies that other styles of helmet will not pass AS/NZS1698:2006, which is not the case. Virtually any style of helmet (i.e., full face, open face, half helmet) can pass AS/NZS1698:2006, as the latter standard has a minimum coverage limit but does not have a "maximum" coverage limit..

This statement is also based on the false premise that AS/NZS1698:2006 prescribes the "style of helmet", which is inaccurate. AS/NZS1698:2006 involves a series of performance specifications and has nothing to do with the "style of helmet".

Figures 2 and 3 in Wordley (2012) show two helmets which fully meet AS/NZS1698:2006: the Bieffe 12R full face helmet[49] (used by DRI in its ATV CPD and ROPS research) and the THH T70 half helmet (also used in Zellner et al. (2014b)).

18. ***"DRI, Quad bike manufacturers and associated industry groups like the FCAI should be extremely careful in publicising the stated benefits of helmet use given the significant differences noted in the types helmets being simulated versus those actually recommended for use." (p 12, paragraph 5)***

As discussed above, this statement is inaccurate and without basis in regard to "the significant differences noted in the type of helmets being simulated versus those actually recommended for use", as the ATV manufacturers do not recommend a particular style or type of helmet, but rather that it meets applicable federal, state, or

---

[48] Anon., "Tips and Practice Guide for the All Terrain Vehicle Rider", page 4, www.atvsafety.org/InfoSheets/ATV_Riding_Tips.pdf, ATV Safety Institute, a Division of the Specialty Vehicle Institute of America.

[49] See Footnote 2.

other standards, which the Bieffe B12 full face helmet analyzed by DRI does meet (i.e., it fully meets AS/NZS1698:2006).

This statement is also now out of date, as the subsequent published paper by Zellner, et al. (2014b) describes results for both the Bieffe B12 full face helmet and the THH T70 half helmet. In terms of normalized injury costs, the net benefit of the rider wearing the half helmet (compared to no helmet) with the Baseline ATV was found to be a statistically significant 49% (36%, 62% confidence interval), whilst for the full face helmet it was a statistically significant 60% (45%, 75% confidence interval). In terms of probability of fatality, the net benefit of the half helmet was found to be a statistically significant 61% (44%, 78% confidence interval), whilst the full face helmet was a statistically significant 75% (55%, 95%[50] confidence interval).

These helmet effectiveness estimates are in the range of helmet net benefits based on ATV accident data (for all types of helmets), as reported by, for example, Rodgers (1990) of the US Consumer Product Safety Commission (CPSC), who estimated that helmet wearing on ATVs reduces the risk of non-fatal head injuries by 64%.

19. "It is possible that DRI's quoted risk/benefit ratios for helmet use potentially over state [sic] the benefits provided by the use of AS/NZS:1698 certified helmet designs." (p12, paragraph 5)

This statement is inaccurate and misconceived, as the Bieffe B12R full face helmet analyzed by DRI *does fully meet* AS/NZS1698:2006.

Simulations done with the THH T70 helmet do result in somewhat lower net benefit (when worn with the baseline ATV) than does the Bieffe helmet, and this is consistent with the former's reduced coverage. It is important to note that the style of helmet does not affect the conclusions regarding the Quadbar, as the Quadbar was found in Zellner, et al (2014b) to have no statistically significant net benefit with either helmet, for the population of overturns. However, the mean (or "point estimate") for the simulation sample indicated that the Quadbar had a more negative net benefit (i.e., had more net harm) with the half helmet than it did with the full face helmet.

4.2.4 Revised AIS Injury Codings

20. "In 2011, shortly before the release of the current paper, they reworked and updated these AIS-codings via an internal technical memorandum, which is referenced by the current work but is not available in the public domain. A copy of this memorandum has been requested … but to date has not been supplied," (p 13, paragraph 1)

[50] There was a transcription error in the Zellner et al. (2014) TRB paper, in that Table 1, Row 3, should say that the 95% confidence interval on probability of fatality net benefit should say "[55%, 95%]" rather than "[35%,115%]". This corrected confidence interval is consistent with the listed p value of p<0.001 and the other statistics listed.

DRI has no record of any such request being made by Wordley, ISCRR or Monash University. The technical memorandum[51] is (and has been since it was issued) available on request from the DRI Library (www.info@dynres.com), and it has also been provided to UNSW researchers, amongst others. If there were more than one request for it, it could be uploaded to the public internet site (www.dri-atv-rops-research.com).

21. ***"DRI have chosen to filter the reported injury data, so that any injury which cannot be predicted by a physical crash dummy is excluded. As a result, the number of injuries translated into AIS codes has been greatly reduced in the updated version, and some inconsistencies are evident." (p 13, paragraph 2)***

This statement could easily give rise to misunderstanding in that it suggests, impliedly, that an inappropriate procedure was somehow involved.

In fact, the "significant changes" to the "reported injury data", i.e., the injury codings, were both required and appropriate because the specific purpose of the subject research was *to enable calibration of the simulation model*, which at the current state of technology is only able to predict a limited set of locations, types and severities of injury (as is the case for all crash dummies applicable to cars, motorcycles, etc., worldwide). It would therefore be *inappropriate, senseless and not possible to attempt to include those injuries that cannot be monitored by a crash dummy*. To include those injuries that cannot be monitored would be to attempt to compare "apples to oranges". This is the reason that some locations, types and severities of injuries were, appropriately, "filtered" (i.e., removed) from consideration.

This statement that "some inconsistencies are present" is unclear in regard to what it is referring.

22. ***"DRI concede that, "in general, current technology (physical) crash dummies cannot monitor for external contusions, abrasions, lacerations, the full range of thoracic and lumbar spinal injury severities, and many other locations, types and severities of human body injury." (p 14, paragraph 1)***

This statement could easily give rise to a misunderstanding that some "concession" has been made by DRI. The passage quoted is not a "concession", but merely a statement of fact.

23. ***"DRI further state that, "Only those actual human injury locations, types and severities that can be monitored by the crash dummy using existing technology can be validly compared." This comment is not strictly correct, as it could be argued that no such comparison between physical crash dummy injuries and reported injuries is being made by this work." (p 14, paragraph 1)***

---

[51]Zellner, et al., 2011.

This statement is incomplete, without clear basis and could easily give rise to misunderstanding, as it is attempting to draw a major – but baseless - distinction between, on the one hand, “physical crash dummy injuries” and, on the other hand, the ATB model of a physical crash dummy from which the injury predictions are simulated.

Certainly, DRI’s ATB model of the ISO 13232 MATD crash dummy has been intended since the beginning to represent and to predict accurately the response of the physical MATD crash dummy, the principle being that in order to be useful, a simulation model must be able to be physically verified, at least in principle, by means of a physical experiment, as a basic principle of the scientific method.

Therefore agreement between the physical dummy and the simulated dummy, in terms of response and injury predictions, should be as close as possible; and to the extent achievable, eliminate any distinction between the physical dummy and the simulated dummy. Accordingly, approximately 33 static and dynamic physical dummy tests and corresponding simulations, and comparison of the response of the physical dummy with the simulated dummy, were reported by Kebschull, et al. (1998) and in Appendix E of Zellner, et al. (2012). Those results indicate extremely close agreement between the physical dummy response (which is the basis for physical dummy injury prediction) and the simulated dummy response (which is the basis for physical dummy injury prediction).

In addition, 12 full-scale motorcycle dynamic impact tests and corresponding simulations, and comparison of the simulated responses and predicted injuries of the MATD dummy against those of the physical MATD dummy, were reported by Kebschull et al. (1998), and those results indicated a high level of agreement in terms of responses and injury predictions between the physical dummy and the same simulated dummy that has been used in the ATV ROPS/CPD research.[52]

In addition, the physical dummy (and ATV) motion response (which lead to predicted injuries) during ATV overturns tests was correlated against those of the simulated dummy, and the Pearson’s correlation coefficient was found to be 0.91, as reported in Appendix F of Zellner, et al. (2012), and, in terms of the Nash-Sutcliffe model efficiency E, E was found to be 0.89, indicating close agreement on that level.

In addition, the injury predictions of the simulated dummy in n=110 type of ATV overturn were compared to actual injuries in n=110 actual overturns (for those injury regions, types and severities that were able to be monitored with a physical/modeled

[52]Although ISO 13232 requires that Pearsons’ correlation coefficient be used as a figure of merit for expressing the level of simulation/test agreement, and was used in Kebschull, et al. (1998), a more appropriate figure of merit would be Nash-Sutcliffe model efficiency (E). The authors have submitted a proposal to ISO Working Group 22 to amend ISO 13232 to replace Pearson’s correlation coefficient with Nash-Sutcliffe model efficiency. For example, the Nash-Sutcliffe model efficiency results quantifying the level of agreement in head resultant accelerations from the simulations and from the tests (in Figure 1 of Kebschull, et al. (1998)) is E=0.89. This numerical value is close to the ideal value of 1.0 and indicates that there is relatively close agreement between the simulation results and the test results.

dummy) in Appendix G of Zellner, et al. (2012). The Nash-Sutcliffe model efficiency E was subsequently determined for this comparison and, for some body regions, the value of E was in excess of 0.99. The average value for E when comparing to the helmeted rider simulations was 0.95. This indicates that, according the Oxford Concise Dictionary definition of “validate”, the simulation was “an adequate representation of the reality it is attempting to model” (i.e., ATV accidents), within the limitations that exist regarding the availability of complete and accurate information on ATV accidents.

Therefore, there is no basis for Wordley (2012) to attempt to draw a major distinction between the physical dummy and the simulated dummy in terms of predicted injuries.

24. ***“Rather the reported injuries are being compared with the injury predictions generated by the ATB simulation package, which conceivably could be modified to include more of these injury modes. DRI themselves have demonstrated that such modifications are possible with the addition of new face fracture, skull fracture and asphyxiation criteria.” (p 14, paragraph 1)***

The above statement by Wordley (2012) is misleading as it suggests that modifications to include additional injury modes were made, readily, by DRI, whereas, in fact, the simulation model “modifications” that were implemented by DRI actually required no modification to the physical dummy and corresponding simulation models in order to monitor specific injury assessment values that have been associated with certain injury modes and for which injury criteria were developed. In particular, the asphyxiation/breathing difficulty criterion can be readily monitored by calibrating the MATD dummy chest deflection (which is currently recorded in the physical dummy) against the static force criterion for “breathing difficulty” (i.e., 490 N (110 lb) (basically, 11 mm of chest deflection corresponds to 490 N (110 lb) of chest compression force). Likewise, a physical dummy skull monitoring system involving peak pressure-recording film was developed and implemented by Smith, et al. (2014) to enable monitoring of the face and skull fracture criteria without modifications to the dummy itself. These injury assessment values can also be calculated and output by the calibrated and validated computer simulation models.

In accordance with ISO 13232 (2005), the current MATD dummy can also monitor for AIS 1 (slight) to AIS 6 (maximal) chest injury due to low velocity chest compression and higher velocity-compression (so-called “viscous criterion”) injuries.

However, the other injuries mentioned by Wordley (2012) (p.15) as being excluded from the updated data set (i.e., shoulder injuries (including fractures, dislocations and lacerations), back injuries, many types of trunk injury (including fractures, contusions, abrasions), collar bone injuries (including breakages and dislocations), arm injuries (upper, lower, elbow, wrist, hand and fingers, including fractures), other site-specific or full body contusions, abrasions, lacerations, bruising and sprains) are currently not monitored by the MATD dummy or any other crash dummy worldwide and, to be monitored, *these injuries would each require major physical modifications to the dummy* due to their relative complexity and the existing structures of the MATD, not to mention substantial research and development of new injury criteria for those body regions. Normally, such modifications are done based on some type of priority or need (i.e., known, significant long term costs to society observed in the given type of accidents with the given type of vehicle). Furthermore, such modifications are usually done in coordination with and by standards committees, and involve multi-national testing, analysis, reporting and peer review, which ordinarily requires years of activity.

25. ***"It is unclear how the arm fractures that DRI identifies are treated, and if these results are included in some form in the results. For example, for CaseID's 11, 96, 100 and 890501BEP0011, DRI report the assumption of radius or ulna fractures (elbow bones) at AIS = 2, but no values are recorded in their results table." (p 14, paragraph 3)***

    This statement is accurate.

    The reason arm injuries are coded in the real world accident tables, but not modeled in the simulation, is that, at an earlier stage of the updated research, consideration was being given to modeling arm fractures, assuming that frangible physical dummy arm bones (lower arm, upper arm) could be designed, fabricated and developed, and arm injury criteria developed, in ways that were analogous to the approach used for the upper and lower legs of the MATD dummy. On further consideration, it was concluded that such physical dummy R&D and modifications would be extensive, and beyond the state of technology. Therefore, arm injury monitoring was not implemented in the simulation model, although the arm injury data remained in the injury coding report.

26. ***"It should also be noted that this update draws upon a reduced total number of incidents (110 rather than 113) following DRI's recent observation that three of the initial cases did not actually involve Quad bikes." (p 15, paragraph 1)***

    This statement requires further clarification.

    The UK/HSE and US/CSPC databases include non-ATVs as well as ATVs, as defined by ANSI/SVIA-1-2010 (2010). As part of the subject research update, DRI reviewed the UK/HSE and US/CPSC cases that were used to develop the 113 ATV overturn types. It was determined during this review that HSE case 25 involved a Kawasaki "KLT 200 198cc TRICKE", which is presumably a three wheel ATV and not relevant to the subject research. It was also determined during this review that HSE cases 24 and 76 involved Yamaha "PRO-HAULER ATV"s, which is a model name that has been associated with side-by-side vehicles and therefore would also not be relevant to the subject research. These three cases were therefore removed from consideration.

Similar errors of improper inclusion of non-ATVs are present in the case descriptions in the Australian NCIS database.[53]

27. "Thus, it is apparent that the following injury regions and types have been excluded from the updated data set:
- Shoulder injuries (including fractures, dislocations and lacerations)
- Back injuries
- Many types of trunk injury (including fractures, contusions, abrasions)
- Collar bone injuries (including breakages and dislocations)
- Arm injuries (upper, lower, elbow, wrist, hand and fingers, including fractures)
- Other site specific or full body contusions, abrasions, lacerations, bruising and sprains." (p 15, paragraph 2)

This statement is accurate, but could easily give rise to misunderstanding as suggesting that the exclusion of these injuries is somehow inappropriate. In fact, as the purpose is to model the physical dummy, and those injuries that the physical dummy can monitor, it would be improper to model the above listed injuries, as they are beyond the state of dummy technology and dummy injury criteria. Developing the capacity to monitor these injuries would involve substantial physical R&D, modification, various types of validation and calibration, not to mention extensive injury criterion development relevant to a physical dummy that can monitor the forces, displacements, etc. that are involved in each of these injury regions and types.

It is essential that, that in order to be useful, a simulation model must be able to be physically verified, at least in principle, by means of a physical experiment; and for valid injury prediction in the vehicle field, this invariably means use (and accurate modeling of) physical crash dummies.

Note that the existing ISO MATD dummy (and corresponding MATD simulation model) can monitor for injuries (from AIS 1 (slight) severity to (in some body regions) AIS 6 (critical/maximal) severity) involving:

- Brain (via 9 accelerometer array)
- Skull/vault (via pressure-recording film)
- Skull/face (via pressure recording film)
- Neck (via 6 axis load cell)
- Chest (via 4 string potentiometers)
- Abdomen (via Rouhana frangible insert)
- Femurs (via frangible bones and strain gauges)
- Knee ligaments (via varusvugus and torsional frangible sheer pins)
- Tibias (via frangible bones and strain gauges)

Many of these involve multi-axis (i.e., frontal, lateral, vertical) sensing. These were judged (by the ISO committees developing and approving the MATD dummy

[53]For example, NCIS, 2009.

specification) to be technically feasible and correlate-able against biomechanical data for cadaveric subjects. These injuries are also those that may be relatively common and/or severe in motorcycle and ATV accidents.

Wordley (2012) seems to imply that unless a dummy can monitor for every type and location of injury, it is not worth monitoring for any injury (which was in fact the approach taken for the Snook (2009) tests described by Wordley (2012)). It should be noted that automotive test dummies also do not monitor every type and location of injury, and might therefore be subject to the same criticism by Wordley (2012). The significant risk of not monitoring for *any* injury types is that false, misleading and baseless conclusions can result, based on speculation and assumptions about what an imaginary rider might be exposed to, as occurred in the Snook (2009) research, as discussed throughout the current report.

28. "The large proportion of recorded injuries being excluded from these body region raises the obvious question as to why DRI did not consider improving and extending their injury simulation of these commonly affected and currently neglected body regions rather than focusing on the addition of yet more head injury mechanisms which were only applied to the non-helmeted condition." (p 16, paragraph 1)

See Items V.21-25 and V.27 above.

This statement is based on two false premises:

1) that it would be trivial to add these other body region injuries (in terms of modifying a physical dummy, providing sensors to monitor for such injuries, and developing new injury criteria for such body regions that could be monitored by the physical dummy (and modifications thereto) and correlated against laboratory test data for cadaveric subjects, which in fact would be an enormous effort, typically involving many years of work; and

2) that inclusion of face and vault injuries somehow had a large effect on the Quadbar risk/benefit outcomes (when in fact it was the significant effect of helmet use upon brain injury that had the larger effect. Furthermore, the addition of the vault and face fracture criteria based on contact force specifically addresses the potential for crushing type injuries associated with opposing contact forces (and therefore would not be associated with large head accelerations that were previously the only head variables monitored) that were considered to be particularly relevant to the subject research.

Overall, ATV serious injuries and fatalities in the published data have been observed to be associated with head, neck and chest injuries (and to a far lesser extent abdominal and lower extremity injuries) and these regions are able to be monitored (i.e., the associated forces and accelerations can be measured with sensors already in, or else readily includable in, the existing MATD crash dummy).

29. ***"The lack of any formally acknowledged arm injury cases and or predictions, and the elimination of approximately reported 25 trunk region injuries (leaving only 14) calls into question the overall validity and generality of the simulation results and the claimed success of what is at best, a heavily filtered correlation with the reported accident injuries." (p 16, paragraph 2)***

    This statement is based on the several false premises discussed in Item V.28 and other items above, and could easily give rise to misunderstanding.

    As discussed above, it is vital to be able to verify/calibrate simulation results by means of physical experiments - in this case, by means of a physically realizable dummy. An enormous volume of R&D, physical modification, validation, calibration and dummy-based injury criteria development would be required to monitor for injuries to these other body regions; and they are not the head, neck and chest body regions that are critical/crucial in ATV fatal/serious injuries, which *are* able to monitored for in the MATD dummy (including skull and face fractures via pressure-recoding film and asphyxiation (via chest potentiometer calibrated in static force units) .

    It would have been *inappropriate* to include these other body regions in attempting to correlate simulation predicted injuries against actual injuries, as no physical dummy exists that can monitor for these other injuries, in order to enable the simulation to be correlated against/verified by physical tests.

    Excluding these other body regions does not ***"[call] into question the overall validity and generality of the simulation"***, as the critical head, neck and chest injuries that are of greatest importance in real ATV accidents are already included in the physical dummy, the simulated dummy and of course, in the real world.

    "Heavily filtered" is an inappropriate term as it ignores the appropriateness of excluding those body regions that cannot currently be monitored using existing technology (including dummy-compatible sensors and injury criteria); and because it ignores the enormous difficulties in monitoring for injuries to these body regions in a physically realizable dummy and development of associated dummy-based injury criteria.

30. ***"Given the relatively short time frame between the release of their updated AIS codings and updated research, and the significant margin for interpretation available, there existed an obvious potential for the injury codings to be pre-emptively tuned to closely match the forth coming results, particularly given the majority of authors are common to both works. Such a bias is not automatically implied, but the opportunity should be acknowledged." (p 16, paragraph 2)***

    This statement is speculative, defamatory, inaccurate and without any basis. Such a process would have been highly inappropriate (not to mention extraordinarily difficult). The codings were simply updated to remove invalid (i.e., non-ATV) cases; to remove body regions and types of injuries that cannot be monitored with existing dummy technology; and to update the codings to the 2005 AIS coding conventions.

Firstly, the fact that arm injuries were left in the injury coding report (as discussed in Item V.25 above), and were not removed based on the ***"forthcoming results"*** of the simulations, is clear evidence that such a ***"pre-emptively tuned"*** process was not undertaken. If the injury codings were ***"pre-emptively tuned"*** based on the results of the simulations, why were the arm injury codings not removed from the report (as arm injuries were not included in the simulation model)?

The actual sequence was not as implied by Wordley (2012), but rather was as follows:

1. Comments were received from several individuals indicating that in the DRI 2004 report and associated data files, the simulation-predicted injuries for the 113 general types of overturn did not closely match the coded actual injuries from the 113 actual cases;
2. DRI, aware that case-by-case matching was not the intention or even technically possible (i.e., only 113 "general types" of ATV overturn could be approximately modeled), investigated additional reasons for the aggregated injury severity distributions not being in closer agreement as between the real accidents and the simulations of the general types of accidents;
3. As a result of that investigation, DRI realized that the simulation-predicted injuries *could not possibly match* the actual injuries coded in DRI-TM-04-77 of December 2004, as the simulation model did not and could not (without enormous efforts and years of R&D to modify and verifiably calibrate the physical dummy in order to monitor for *all* of the injuries that had been coded;
4. Therefore, DRI focused its attention on those injuries that *could be monitored* with the physical MATD dummy (including relatively small and relatively simple extensions of the MATD monitoring/injury criteria for asphyxia, skull and face fractures);
5. After the injuries that were not monitor-able by the physical (and therefore the simulated) dummy were removed from consideration, there was still not as strong a correlation between simulated and real aggregated injury severity distributions as would be desired (i.e., as described in Zellner, et al. (2012), there was some over-prediction of head and leg injuries). Therefore, attention was focused on checking and considering improvements to the fine details of the simulation model that affected those particular injury outcomes, and performing additional measurements and improving modeling and estimation methods for a long list of these (e.g., head/helmet force-deflection characteristics; leg bone force-deflection characteristics; head/ground, helmet/ground and dummy/ground friction characteristics; and the "low energy" assumption that an overturn barely occurred, etc.) were systematically investigated and implemented;

6. The results after making all of these refinements were that, with the baseline ATV and n=110 overturn types, there was a much higher agreement (i.e., Nash-Sutcliffe model efficiency (E=0.95)) in terms of predicted and actual aggregated injury severity distributions than had previously been the case;
7. Then and lastly, the Quadbar simulations were run, and the resulting injury indices were compared to the Baseline injury indices, in order to conduct the risk/benefit analysis of the Quadbar.

4.2.5 Methods used to generate unknown scenario variables

31. ***"The final unknown values used by the simulation were found through a "systematic variation" procedure which is illustrated with system diagrams." (p 17, paragraph 1)***

    This statement could easily give rise to a misunderstanding that DRI was attempting to identify unknown initial condition values in the original accident case, which was not what was being done. In fact, as a clarification, for those initial conditions that were not described in the original accident case, the "systematic variation" procedure involved making consistent and plausible assumptions for those unknown variables <u>for the entire category</u>, or general type, of overturn, within which the original actual accident would lie, such that an overturn would have barely occurred. This process and the results of applying it are described in Zellner, et al. (2012) Appendices C and D.

32. ***"These new incident interpretations have been called "low energy" overturns, in comparison to the previous overturns which were classified as "high energy", based on the assumption that they used comparatively more extreme values to guarantee an overturn." (p 17, paragraph 1)***

    This statement is unclear, as it was in fact more than an "assumption". Zellner, et al. (2012) Appendix B, "Original methodology used to define the 113 US/UK ATV overturn types", contains a more severe set of assumptions than Appendix C – "Updated "low energy" methodology used to define the 110 updated US/UK ATV overturn types". The less severe (lower energy) nature of the Appendix C assumptions is due to gradually increasing the severity of the initial conditions by relatively small increments until overturn occurs, rather than defining arbitrarily severe initial conditions *a priori* as in Appendix B.

    By definition, the maximum elevation change (of 3 metres) was generally substantially lower than in the original simulations, which had arbitrarily long slopes.

    As stated in Zellner, et al. (2012), and with all other factors being equal, "By way of comparison, in the original "high energy" overturn types, the dummy's final kinetic energy (at the end of the simulation run) was less than 1% of its initial kinetic energy (i.e., it was virtually at rest) in 59% of the baseline ATV overturns; whereas in the "low energy" overturn types, the dummy's final kinetic energy was less than 1% of its initial kinetic energy in 99% of the baseline ATV overturns." This verifies that the updated initial conditions were indeed less severe.

33. "These new low energy overturns reportedly resulted in the need for much longer simulation times, and consequently each simulated run was solved for a total of 10 seconds rather than the 6 seconds used previously." (p 17, paragraph 1)

This statement is inaccurate, as the "low energy" overturns generally resulted in the dummy coming to rest sooner, rather than later. Instead, the longer (i.e., 10 s) simulation run times were an effort to allow the dummy to come to rest completely (which was not the case in all of the previous Munoz et al. (2017) simulations which had ended at 6 seconds), particularly so as to enable potential entrapment and asphyxiation/breathing difficulty index to be evaluated (which had not been evaluated in the previous simulations, due to there being no asphyxiation criteria available at that time). In the previous "high energy" simulations, the vast majority of non-asphyxiation-type, non-skull-crush-type injuries had occurred within 6 seconds, so at that time it was judged not to be essential that the dummy came to rest. This is explained in Zellner, et al. (2012), page 60.

34. "It may be that inaction on the part of the static rider could render such simulated accidents highly unrealistic and unrepresentative of "what a real person would do". A review of the simulation video outputs is required to comment in more detail on this potential issue. Given that DRI have claimed that passive riders are still the "state of the art", and that it is not currently possible to simulate more realistic human behaviour, serious consideration should be given to the question of whether any of these low energy, low speed, long duration simulation results can be trusted at a very fundamental level. In this specific instance, the best computer science can currently do may not be good enough, which invites us to investigate and develop more appropriate techniques for quantifying the performance of potential safety devices." (p 17, paragraph 1)

This statement is unclear and inaccurate, as the dummy is not "static", but rather is "passive", dynamically responding to dynamic forces during each overturn.

In addition, this statement has no apparent basis other than speculation.

Van Ee, et al. (2012) reports that:

- 28% of n=129 actual riders analyzed on overturning ATVs (from internet videos) *did not attempt active dismount* from the ATV (while 72% attempted active dismount, and 45% successfully dismounted);
- Those that did not attempt active dismount from the ATV had more than twice (i.e., 32%) the injury rate of those that did attempt dismount (i.e., 15% injury rate).[54]

[54]Van Ee, et al. (2012), page 5.

– "Active rider dismount was a common and effective strategy to avoid injurious ATV-rider contact. Video analysis and laboratory investigation demonstrates that a rear mounted CPD may obstruct rider dismount and successful separation."[55]

These results indicate that riders "not attempting dismount" are more "at risk" of injury than those attempting active dismount. This supports the appropriateness of simulating such passive (i.e., "not attempting dismount") riders, who are the population most at risk.

Likewise, this suggests the hypothesis that a substantial fraction (i.e., up to more than ¼ of riders) involved in these actual overturns were relatively "passive", in effect, they may have been responding in a way that is similar to passive crash dummies. In order to test this hypothesis, it should be possible in principle to digitize the Van Ee (2012) videos, run simulations of those overturn scenarios using a passive dummy, and correlate the motions of those riders "not attempting dismount" with those of the passive dummy. For those riders for which a strong correlation exists, this test would demonstrate that at least some riders (and potentially as many as ¼ of riders) involved in actual overturns -- and those who are at higher risk of injury -- respond, in effect, in a way which is similar to passive crash dummies.

This statement by Wordley (2012), in addition to speculating without clear basis about the relevance of passive crash dummies, does not propose an alternative method for evaluating the effectiveness (and any negative unintended consequences) of CPDs in regard to injuries during overturns. Since experiments with live humans in potentially injurious vehicle overturns (be they with cars, trucks or ATVs) presents obvious major ethical issues, it is not clear what Wordley (2012) is proposing as an evaluation method, except possibly the "imaginary dummy" methods used by Snook (2009), which Wordley (2012) appears to support, which has major deficiencies, as discussed in Items II.2, II.12 through II.14, III.7 through III.9, and more extensively in Zellner, et al. (2012b).

See Item V.12 regarding the above statement in Wordley (2012) regarding "review of simulation video outputs".

4.2.6 Process used to generate additional scenario perturbations

35. *"These variations are made in a set order: steering, followed by braking, speed, obstacle height, initial heading and slope, up until the point that an additional six cases, all resulting in roll over, have been generated. It is unclear whether this process was used (but not reported) in the initial work, or if this methodology is a new feature of the updated research." (p 17, paragraph 2)*

This statement is inaccurate in regard to any claimed lack of clarity, as both the original perturbation method (first used in the Zellner, et al. (2008) FISITA paper) and the updated perturbation method are described in Zellner, et al. (2012).[56]

[55]Op cit., page 2.

As discussed in Item V.III and Footnote 21 herein, the first time this perturbation (i.e., sensitivity analysis) method was used, in the Zellner, et al. (2008) FISITA paper; the methodology was described in Table 4 of the draft of the FISITA paper; the FISITA paper reviewers requested that the paper be substantially shortened to meet their length criteria, resulting in removal of Table 4; and the next opportunity to describe the original perturbation method was in Table 4 of Zellner, et al. (2012).

36. "Again, to properly assess the changes made to the generation of unknown scenario variables and the scenario perturbations it is necessary to carefully review the video outputs, at least for a representative range of the 3080 individual simulations conducted." (p 18, paragraph 2)

See Item V.12 regarding the above statement in Wordley (2012) regarding "review the video outputs".

4.2.7 Changes to the baseline(s) used to normalise reported results

37. "The many modifications that DRI have made to their modelling and simulation techniques are over shadowed and potentially misrepresented by a critical change they have simultaneously made to the way in which they present their comparative results." (p 18, paragraph 2)

This statement could easily give rise to misunderstanding in that it inaccurately suggests that ***"the way in which [DRI] present their comparative results"*** somehow ***"overshadows"*** the modifications made, potentially ***"misrepresents"*** the results, and is a ***"critical change"***.

First, it is not a "critical change", as the updated results presented in Appendix I of Zellner, et al. (2012), which is based on DRI's previous reporting method, indicate the Quadbar has no statistically significant net benefit, and its risk/benefit is not statically significantly different from 100%, which is much greater than the risk/benefit for any known vehicle safety device, which is *very similar to the overall result* reported in Munoz, et al. (2008). Wordley (2012) elsewhere confirms there is ***"very little change"*** in the overall results using the "two baseline" method,[57] but the above statement describes the new "single baseline" reporting method as a "critical change".

Second, the "single baseline" results are not "misrepresented", as the results reported in the main text of Zellner, et al. (2012) clearly state that they are for the "baseline

---

[56]The original perturbation method is described in Zellner, et al. (2012), Table 4, and the updated perturbation methods is described in Zellner et al. (2012), Appendix K.

[57]Wordley (2012), page 20, "The modified test methodologies and models used in the updated 2012 research have resulted in very little change to the overall risk/benefit results, based on the more appropriate multiple baseline comparison [as reported in Zellner, et al. (2012) Appendix I]."

ATV with helmet"; and as noted above, the main text cross references the results using the previous "two baseline" method.[58]

Third, the "single baseline" method is the more appropriate method to include in the main text because the "two-baseline" approach can be misleading in that it shifts focus away from a primary hazard, which is (in this case) not wearing a helmet. Not wearing a helmet is approximately 2.5 times more hazardous than wearing a helmet, and even if a device such as a CPD reduced that hazard by a small amount – which the DRI results for the Quadbar do not indicate – the unhelmeted hazard – and conversely the large benefits of wearing a helmet – *far exceed* the estimated effects of fitting a CPD (while unhelmeted), *which were found to be statistically insignificant*.

Fourth, the "single baseline" method is the more appropriate method to include in the main text because the "single baseline" results (presented in the main text of Zellner, et al. (2012)) can be considered to be a *combination* of the (previously used) two separate "two baseline" results (presented in Appendix I of Zellner, et al. (2012)). Therefore the "combined" result is appropriately reported in the main text, and the two "component" results are appropriately reported in Appendix I. This combined effect is further described in Zellner et al. (2014a).

Fifth, other reasons why the "single baseline" method is the more appropriate method to use in the main text are enumerated in the main text[59] and in Appendix B of Zellner, et al. (2012).

Sixth, as noted therein, the "two baseline" method, which was previously the only method used, can be potentially misleading (and consequently harmful) to the public. For example, US first generation airbags were originally only regulated, tested and tuned so as to be a "net benefit" (and a low risk/benefit) for one "misuse" condition (i.e., unbelted mid-size adult males) whereas it was later discovered that such over-tuning for this baseline resulted in a high risk/benefit percentage for another (i.e., alternate baseline) condition (i.e., small front seat passengers), resulting in approximately 380 airbag-induced fatalities to the latter category as of 1998.[60] As stated elsewhere, using a "single baseline" references the results to a single (ideally the intended) use, and avoids misinterpretation of such confusing and contradictory "multiple baseline" outcomes. This is also relevant to the ATV CPD case, as discussed immediately above, as the effect of wearing a helmet (i.e., first baseline) is much larger than and overwhelms any effect of the CPD (second baseline), an outcome which is not evident from the "two baseline" results.

Seventh, the single baseline method, though not the only method reported, expresses what may occur if riders who currently wear helmets mistakenly remove them under the mistaken assumption that with a CPD, helmets are not needed. This is a scenario that may occur and is therefore relevant.

[58]Zellner, et al. (2012), page 64.
[59]Op. cit.
[60]Iijima, 1998.

Eighth, this statement by Wordley (2012) is now out of date, as Zellner, et al. (2014a) and Zellner, et al. (2014b) describe both the single baseline outcomes and the two baseline outcomes in the main conclusions.

38. ***"Due to the significant proportion of recorded injuries and fatalities where the victims were not wearing safety helmets, it is common practice, and indeed common sense, to state the injury risks or benefits in terms of a fair comparison between riders either with helmets, or without them." (p 18, paragraph 3)***

First, this statement is inaccurate because the "two baseline" is not "common practice": there are no international standards or other standards indicating it should be done, ISO 13232 (2005) does not use it, and there is no original research other than DRI's that uses the "two baseline" method that Wordley (2012) prefers.

Second, this statement can be easily misinterpreted because so-called "common sense" conclusions based on a "two baseline method" can easily misinterpreted, as indicated in the sixth point in Item V. 37.

Third, the single baseline method, though not the only method reported, expresses what may occur if riders who currently wear helmets mistakenly remove them under the mistaken assumption that, with a CPD, helmets are not needed. This is an important scenario that may occur and is therefore relevant.

Fourth, this statement is irrelevant on the broadest level, as the "two baseline" results preferred by Wordley (2012) indicate that the Quadbar has *no statistically significant net benefit*, and that its risk/benefit is not statistically significantly different from 100%, which is much greater than the risk/benefit for any known vehicle safety device.

Fifth, this statement by Wordley (2012) is now out of date, as Zellner, et al. (2014a) and Zellner, et al. (2014b) describe both the single baseline outcomes and the two baseline outcomes in the main conclusions.

See Item V.37 for other related points.

39. ***"Despite this observation, in the presentation of their new and updated results pertaining to the effectiveness of the Quadbar CPD, DRI have elected to use "a single baseline condition (i.e., the baseline ATV with helmet, which represents the intended use of the vehicle) in describing the main results." (p18, paragraph 5)***

This statement is based on a faulty premise, as the "observation" being referred to (i.e., that only a two baseline method is common practice and common sense) is not valid, as discussed in Items V.37 and 38.

In addition, the single baseline method, though not the only method reported, expresses what may occur if riders who currently wear helmets mistakenly remove them under the mistaken assumption that, with a CPD, helmets are not needed. This is a possible scenario that may occur and is therefore relevant.

In addition, this statement is incomplete, as it does not mention that DRI also presented (i.e., reported) discussed and cross referenced from the main text the "two baseline" results in Appendix I of Zellner, et al. (2012).

In addition, this statement by Wordley (2012) is now out of date, as Zellner, et al. (2014a) and Zellner, et al. (2014b) describe both the single baseline outcomes and the two baseline outcomes in the main conclusions.

40. ***"It can be seen that the first calculation B1 is no different to A1, however the change between A2 and B2 results in an unfair and potentially misleading comparison being drawn, where the effectiveness of the standard bike with ROPS/CPD but without a helmet is measured against the new "singular" baseline of the standard bike with no ROPS/CPDs but with a helmet." (p 19, paragraph 1)***

This statement is inaccurate as the "single baseline" result is not "an unfair and potentially misleading comparison". To the contrary, as discussed in Item V. 37 through 39, the "two baseline" method preferred by Wordley (2012) and properly presented by DRI in Appendix I can be "an unfair and potentially misleading comparison". It can in principle easily result (and might have resulted except for the fact that the results were statistically insignificant) in the false conclusion that a device is desirable if it has a somewhat greater reduction in risk/benefit for one condition, than it has an increase in risk/benefit in another condition, whereas the overwhelming hazard of the first condition is totally invisible in the results.

The "single baseline" result is the "combination" of the "two baseline" component results, as discussed in Item V. 37 above and as further discussed in Zellner, et al. (2014a), and it is therefore a fair and appropriate comparison.

In addition, the single baseline method, though not the only method reported, expresses what may occur if riders who currently wear helmets mistakenly remove them under the mistaken assumption that with a CPD, helmets are not needed. This is a scenario that may occur and is therefore relevant.

In addition, this statement by Wordley (2012) is now out of date, as Zellner, et al. (2014a) and Zellner, et al. (2014b) describe both the single baseline outcomes and the two baseline outcomes in the main conclusions.

41. ***"With regard to comments c) and d), it seems more logical to assume that multiple and appropriate baselines would be more likely to create potential for correct understanding and interpretation of these results, and the generation of valid***

comparisons of the effects of the addition of ROPS/CPDs for both recommended use (helmeted) and foreseeable use (unhelmeted) conditions." (p 19, paragraph 4)

This statement is biased in terms of what the "correct understanding" is. A more complete view is that *both methods* should be used, in order to provide a more *complete* understanding of the outcomes. That is in fact what has been done more recently in Zellner, et al. (2014a) and Zellner, et al. (2014b).

As discussed in Section V.37 through 40, use of only the two baseline method can result in the potential misunderstanding of appearing to recommend that riders not wear helmets and fit a Quadbar in order to somehow benefit (from the statistically insignificant benefit) of that configuration. First of all, the benefit is statistically insignificant, which means that there is a substantial probability that it is a net harm (as discussed in Section V.44).

Second, the data indicated that wearing a helmet is 2.5 times safer than not wearing a helmet, and therefore that effect overwhelms the other results. Therefore, if the goal is to increase safety, the helmeted case should be given much higher priority in terms of configuration.

This statement by Wordley (2012) is now out of date, as Zellner, et al. (2014a) and Zellner, et al. (2014b) describe both the single baseline outcomes and the two baseline outcomes in the main conclusions.

42. ***"A cynical read of these changes might observe that casual readers are being invited to attribute the drastic changes in the updated results to improvements in the capabilities and accuracy of the modelling, rather than the changes which have been made to commonly accepted comparison practices." (19, paragraph 5)***

This statement is inaccurate, as DRI reported results for *both* methods (in the main text and in Appendix I, the main text cross-referencing the latter), and as discussed elsewhere herein, there are no "commonly accepted" comparison practices in this field.

43. ***"The single baseline method results in dramatic and potentially misleading increases (428% and 274%) in the reported injury risks due to the implied use of the CPD." (p20, 1st bullet)***

This statement is inaccurate, as the "single baseline" method is not "potentially misleading" to anyone who reads the report, since, as discussed elsewhere in this report, the Zellner, et al. (2012a) results clearly state *what* they are in comparison to: that is, the main text explains that the "single baseline" results are in comparison to the "helmeted baseline ATV", cross-referencing Appendix I for the two baseline results; and Appendix I explains that the "two baseline" results presented therein are in comparison to the "helmeted" and the "unhelmeted" baseline ATV.

This statement by Wordley (2012) is now out of date, as Zellner, et al. (2014a) and Zellner, et al. (2014b) describe both the single baseline outcomes and the two baseline outcomes in the main conclusions.

Both single baseline results and two baseline results were reported, as they express the different relative changes in risk/benefit and net benefit that may occur, depending on different starting points.

The single baseline method, though not the only method reported, expresses what may occur if riders who currently wear helmets remove them under the mistaken assumption that, with a CPD, helmets are not needed. This is a important scenario that may occur and is therefore relevant.

For the two baseline method (CPD with no helmet to Baseline with no helmet; and CPD with helmet to Baseline with helmet), as in previous reports, it is clearly stated that the Quadbar had no statistically significant net benefits. As further discussed in Zellner, et al. (2014b) and in Section V.4 above, the associated p-values for these non-significant findings indicates that there is in fact a substantial probability that the Quadbar has a net harmful effect.[61] As stated in Zellner, et al (2014b) "In the presence of such large uncertainties as to the direction of the outcome (i.e., harmful or beneficial) stemming from large p-values (e.g., p=0.202), it is traditionally considered not advisable to proceed with implementing such a treatment (i.e., in this case, the CPD with no helmet, instead of no CPD with no helmet)."

44. ***"For the helmeted condition the risk/benefit prediction has increased slightly by 7%, from 99% to 108% (predicted slightly less safe), while for the no helmet condition the risk/benefit prediction has decreased slightly by 3% (slightly more safe), from 71% to 68%." (p20, 2nd bullet)***

    This statement is inaccurate, incomplete and potentially misleading because reporting these "point estimates" (99%, 108%, etc.) without the corresponding 95% confidence intervals or p-values is incomplete and potentially misleading, and therefore meaningless. Properly stated with the 95% confidence intervals, it becomes clear that the 99% [53%, 192%] and 108% [69%, 168%] risk/benefit estimates are statistically indistinguishable from each other and from 100%, i.e., the injury risks are equal to the injury benefits, for the population of overturns. Likewise, the 71% [41%, 135%] and 68% [42%, 114%] risk/benefit estimates are statistically indistinguishable from each other and from 100%, i.e., the injury risks are equal to the injury benefits, for the population of overturns. In addition, as discussed in Zellner, et al. (2014b) and in Sections V.4 and V.43 above, the large p-values associated with all of these findings indicate that there is a substantial probability that the Quadbar has a net harmful effect.

[61] For example, as stated in Zellner, et al. (2014b), "The associated p-value of 0.202 for the probability of fatality outcome can be interpreted as meaning that the probability that the actual net benefit is *negative* (i.e., that it is harmful in terms of fatalities) is 0.101."

This statement is inaccurate, incomplete and potentially misleading because the terms "predicted slightly less safe" and "slightly more safe" (and the associated 108% and 68% values) *only* apply to the *DRI simulation sample*. In addition, a 108% risk/benefit percentage indicates greater harm than benefit, which at best is "slightly unsafe" rather than "slightly less safe" as claimed by Wordley (2012). To be accurate, this sentence should say "predicted slightly unsafe assuming the *DRI simulation sample are the only overturn types that ever occur*" and "slightly more safe assuming the *DRI simulation sample are the only overturn types that ever occur and that riders should remove their helmets*". As clearly stated in the DRI report, the results for *all samples*, i.e., the "population estimates" are that the effects of the Quadbar are *statistically insignificant*, for both the helmeted and the unhelmeted condition.

Wordley (2012) cannot have it both ways: Wordley (2012) cannot simultaneously *reject* DRI's simulation sample as not being valid, which is stated and implied at various places in the Wordley (2012) report, and at the same time *accept the specific results* that apply *only* to DRI's simulation sample, and not to the population of all overturns.

As noted, the overall results for the population of overturns (as opposed to the results for the *DRI simulation sample*) reported by DRI are that the effects of the Quadbar are *statistically insignificant*, within a 95% confidence interval (equivalent to p=0.05, Box, et al. (1978) is by far the most widely accepted standard for being "somewhat convinced" of the reality of an outcome), for both the helmeted and the unhelmeted conditions. The large p-values associated with all of these findings indicate that there is a substantial probability that the Quadbar has a net harmful effect for the population of overturns.

45. "Given DRI's stated 95% confidence limits for the updated, multiple baseline results (i.e. helmet: [69%, 168%], and no helmet: [42%, 114%]) it is evident that both these results still straddle the neutral risk/benefit value (100%). Hence according to both DRI and the recommendations made by ISO 13232 these results should be considered statistically insignificant. It should be noted though that for the no helmet condition, the addition of the QuadBar CPD is very close (within 14%) of providing what DRI would classify as a statically significant increase in safety (95% confidence, 5% likelihood of error). By comparison, the confidence limits for the helmeted case straddle the neutral value almost symmetrically." (p21, paragraph 2)

It is unclear what ***"within 14%"*** is actually referring to. If ***"within 14%"*** is referring to upper bounds of the "68% [42%, 114%]" estimated risk/benefit ratio for the unhelmeted two baseline comparison, then the statement that ***"the addition of the Quadbar is very close (within 14%) of providing what would classify as a statistically significant increase in safety (95% confidence, 5% likelihood of error)"*** is inaccurate, confusing, and potentially misleading.

The 68% [42%, 114%] estimated risk/benefit result corresponds to an estimated 12%±17% net benefit, where the 12% is the point estimate for the net benefit and the ±17% is the relative 95% confidence interval for the point estimate (i.e., the simulation sample mean) corresponding to a two-sided significance test. The upper limit for the risk/benefit (i.e. 114%) corresponds to the lower limit for the net benefit (i.e., 12%-17%=-5% (which is *a net harm*)) and vice versa. Therefore if the true but unknown net benefit is -5% (which is *a net harm*) and the corresponding risk/benefit ratio is 114% (i.e., injury risks 14% greater than injury benefits), then the probability that a net benefit value of 12% or greater (i.e., more extreme difference) would have been observed is 0.025 (which is small). This means we can be "somewhat convinced" that the risk/benefit percentage is not greater than 114% (which would be a substantially harmful device). However, these same results also indicate that risk/benefit percentages between 100% and 114% (indicating a net harm) are certainly possible. This categorically *does not mean* that the 114% risk/benefit is ***"very close … [to] a statically significant increase in safety".*** In fact, as discussed in Section V.9, the fact that the probability value (p=0.153) for the Quadbar unhelmeted risk/benefit outcome is *more than three times* the "conventional "critical" significance level in common use" (i.e., p=0.05, Box, et al (1978)) indicates that the Quadbar was *not* "very close" to returning a statistically significant prediction of increased safety. Instead it means that there is a 7.7% probability that the Quadbar/unhelmeted in fact has a net harmful effect for the population of overturns, with a much greater probability of a net harmful effect for the Quadbar/helmeted case (i.e., 62%, Sections V.45-47, VI.9).

The subsequent Zellner, et al. (2014b) Table 1 results indicate that the associated p-value for the unhelmeted case is p=0.153. This indicates that if the true but unknown net benefit of the device is 0% and the risk/benefit ratio is 100% (i.e., the Null Hypothesis), then the probability that a net benefit value greater than or equal to 12% would have been observed is p/2=0.0765 and the probability that a value less than or equal to -12% is also observed is p/2=0.0765. Therefore since the p-value is greater than 0.153 we *cannot reject* the Null Hypothesis (i.e., that there is no net benefit) that is based on these results; and therefore the net benefit of the device is considered not statistically significant. Treatments having net benefit p-values substantially exceeding 0.05 (whether they be a pharmaceutical, a medical device, a medical procedure or a safety intervention) are most often not released for clinical trials on humans, because one cannot be "somewhat convinced" that any net benefit suggested by that sample is real, or in fact that the treatment is not harmful (i.e., having a negative net benefit).

In addition, the statement that "the confidence limits for the helmeted case straddle the neutral value almost symmetrically" is unclear in its meaning and implication and could easily be misunderstood as implying that somehow the result is "neutral". For example, the "helmeted case" (i.e., the Quadbar/helmeted in comparison to the Baseline ATV/helmeted, also described in Zellner, et al. (2014b) Table 1) indicates that the net benefit for the normalized probable injury cost is -3% (p=0.760). Therefore, the probability that the true net benefit is less than 0 (i.e., harmful) is 1-

p/2=0.620 (i.e., greater than 50% chance from a 50/50 coin toss). According to widely accepted practice, it is not reasonable, prudent or ethical to recommend a device with this much potential to worsen the injury outcome

4.2.9 Real world context in interpretation of the helmeted versus non-helmeted results

46. "Such a statement appears specifically designed to dissuade anyone from seeing any value in the previous observation that the QuadBar CPD is very close to providing a statistically significant injury benefit, even by DRI's own data." (p21, paragraph 5)

This statement is based on the false premise that the "Quadbar CPD is very close to providing a statistically significant benefit". As discussed in Section V.45, DRI found that there was a 0.077 probability that the Quadbar with an unhelmeted dummy was actually harmful compared to the Baseline ATV with an unhelmeted dummy, and that there was a 0.620 probability that the Quadbar with a helmeted dummy was actually harmful compared to the Baseline ATV with a helmeted dummy. According to widely accepted practice, it is not reasonable, prudent or ethical to recommend a device with this much potential to worsen the injury outcome.

47. "It would be counter intuitive to dismiss the predicted safety improvements due to the use of CPDs for these scenarios on the basis that such scenarios do not conveniently represent the manufacturers' "intended use." (p 21, paragraph 6)

This statement is inaccurate and could easily give rise to misunderstanding as meaning that there were "predicted safety improvements". First, the Quadbar had no statistically significant predicted net safety benefits. The Zellner, et al. (2012a) results that are being discussed here were that there were no statistically significant net safety benefits for either the Quadbar/unhelmeted (compared to the Baseline ATV/unhelmeted) or the Quadbar/helmeted (compared to the Baseline ATV/helmeted) configurations.

In addition, as discussed in Sections V.45 and V.46 above, the estimated probability that the Quadbar has a net harmful effect for the population of overturns is 0.077 for the unhelmeted (two baseline) case, and 0.620 for the helmeted case. It is not reasonable, prudent or ethical to recommend a device with this much potential to worsen the injury outcome

Moreover the "intended use" of an ATV is with a helmet, which is stated on the mandatory warning labels on all ANSI/SVIA 1 (2010)-complaint ATVs. As indicated in the simulation results, as well as in real world accident data (e.g., Rodgers (1990)), wearing a helmet on an ATV is approximately 2.5 times safer than not wearing a helmet..

If one accepts the specific DRI simulation sample as a valid sample of the population, then the simulation result for the population estimate is that the Quadbar has no statistically significant net benefit, for either the helmeted or the unhelmeted configuration.

If one rejects the specific DRI simulation sample as a valid sample of the population and considers that only the overturns in the specific DRI simulation sample and the point estimates occur, then it is not reasonable, prudent or ethical to recommend a device (i.e., the Quadbar) that worsens the injury outcome for those who wear helmets (i.e., for whom the injury risks of the Quadbar are greater than the injury benefits, by a ratio of 108%).

48. "As highlighted previously, researchers such as DRI, the FCAI and Quad bike manufacturers also need to clarify which style, or styles, of helmet they are simulating and which they are recommending. It is not possible to assume or claim that AS/NZS:1698 compliant helmets that are broadly recommended can provide the same injury protection as the more substantial, full face helmets that DRI are currently simulating.".(p 21, paragraph 6)

This statement is inaccurate, based on several false premises and is liable to give rise to misunderstanding.

In addition, this statement is now out-of-date as Zellner, et al. (2014b) conducted analyses with both a full-face helmet and a half helmet (both of which fully comply with AS/NZS:1698) and the conclusions regarding the Quadbar are not affected in any way by which style of helmet is used.

The first false premise is that DRI did not clarify which style of helmet is being simulated. This is inaccurate, as Zellner, et al (2012a) states that a "full-face helmet such as the one specified in ISO 13232 (2005)"was used.

The second false premise is that FCAI and the Quadbike manufacturers have not been clarifying ***"which styles, or styles, of helmet…they are recommending"***. As discussed in Section V.16, the ATV Safety Institute, with which all manufacturers' helmet recommendations are in agreement, recommends wearing a helmet that complies with applicable state standards, and recommends both full-face and open faced helmets (with added mouth protection). Likewise, FCAI recommends that helmets comply with AS/NZS:1698 (with which both full face helmets and half helmets can comply) and does not recommend any particular style of helmet.

The third false premise is that someone, or anyone, is claiming that ***"AS/NZS:1698 compliant helmets that are broadly recommended can provide the same injury protection as the more substantial, full face helmets that DRI are currently simulating."*** First of all, the Bieffe B12 full face helmet that DRI simulated *does* comply with AS/NZS:1698.[62] Secondly, no one has claimed that a half helmet provides the same level of protection as a full face helmet. As indicated in Section V.16, the ATV Safety Institute website states that "Full-face helmets help protect

[62] Smith, T.A, Koerner, J.E., 2014.

your face as well as your head. Open-face helmets are lighter and may be cooler, but should be used with mouth protection. Eye protection should be used with both types of helmets." This indicates advantages of both styles of helmet, as well as different levels of protection (i.e., of the face). Zellner, et al. (2014b) found that the Bieffe B12 full face helmet had a net injury benefit of 60% (45%, 75% confidence interval) while the THH T70 half helmet had a 49% (36%, 62% confidence interval) indicating different levels of protection for these two specific helmets but did not attempt to generalize this to their respective helmet types.

49. ***"The second risk/benefit estimate for the unhelmeted condition (1088%) is however invalid, due to the inappropriate baseline comparison used (i.e. no CPD, but with Helmet), and should be disregarded." (p 22, paragraph 2)***

This statement is inaccurate, as the ***"second risk/benefit estimate for the unhelmeted conditions (1088%)"*** is not ***"invalid due to the inappropriate baseline comparison used"***. The "intended use" (i.e., helmeted) baseline is entirely appropriate and valid as it is the recommended and intended condition of ATV use, against which various countermeasures can be compared. One of these countermeasures is the CPD with helmet and another is CPD without helmet. As discussed in Zellner, et al. (2012a), pp 63-64, Sections II.8 through 10, IV.7, and V.37 through 43 above, the single baseline method is relevant for a list of reasons, including that (a) it is relative to the intended (safer) condition (i.e., helmeted) and (b) it can capture what occurs if a user were to falsely assume that the CPD has a significant protective effect and that he/she no longer needs to wear a helmet.

50. ***"Without provision of this [fatality probability] result or the underlying raw data, it is not possible to determine if the CPD would have provided a statistically significant fatality benefit for the true unhelmeted comparison." (p23, paragraph 1)***

The probability of fatality data were inadvertently omitted from the Zellner, et al. (2012) report, and subsequent publications of these same analyses (along with extended analyses) have included the probability of fatality data (i.e., in Zellner, et al. (2014a) and Zellner et al. (2014b)). For example Table 1 in Zellner, et al. (2014b) applies to probability of fatality and is as follows:

Table 1. Summary of Risk/Benefit and Net Benefit Percentage Results

Baseline	Counter-measure	Injury Index	Estimated Risk/Benefit Percentage	Estimated Net Benefit	Statistical Significance
No helmet, No CPD	Full face helmet	ICnorm	**9% (6%, 21%)**	**60% (45%, 75%)**	**Significant $p<0.001$**

Baseline	Counter-measure	Injury Index	Estimated Risk/Benefit Percentage	Estimated Net Benefit	Statistical Significance
	No CPD	Prob. of Fatality	**2% (1%, 18%)**	**75% (55%, 95%)**[63]	**Significant p<0.001**
	Half Helmet	ICnorm	**13% (9%,25%)**	**49% (36%, 62%)**	**Significant p<0.001**
	No CPD	Prob. of Fatality	**6% (4%, 23%)**	**61% (44%, 78%)**	**Significant p<0.001**
	No helmet with CPD	ICnorm	68% (42%, 114%)	12% (-5%, 29%)	**Not Significant** p=0.153
		Prob. of Fatality	68% (41%,120%)	14% (-7%, 35%)	**Not Significant.** p=0.202
Full face helmet, No CPD (intended use)	Full face helmet and CPD	ICnorm	108% (69%, 168%)	-3% (-24%, 18%)	**Not Significant.** p=0.760
		Prob. of Fatality	134% (79%, 219%)	-15% (-44%, 14%)	**Not Significant.** p=0.309
	No helmet with CPD	ICnorm	**492% (255%,788%)**	**-114% (-154%, -74%)**	**Significant p<0.001**
		Prob. of Fatality	**1088% (322%,1987%)**	**-244% (-327%,-161%)**	**Significant p<0.001**
Half helmet, No CPD (intended use)	Half helmet and CPD	ICnorm	153% (96%, 236%)	-20% (-41%, 1%)	**Not Significant.** p=0.074
		Prob. of Fatality	158% (89%, 265%)	-24% (-55%, 7%)	**Not Significant.** p=0.132
	No helmet with CPD	ICnorm	**348% (193%, 548%)**	**-70% (-100%,-40%)**	**Significant p=<0.001**
		Prob. of Fatality	**504% (215%, 822%)**	**-122% (-173%,-71%)**	**Significant p=<0.001**

[63] There was a transcription error in the Zellner et al. (2014) TRB paper, in that Table 1, Row 3, should say that the 95% confidence interval on probability of fatality net benefit should say "[55%, 95%]" rather than "[35%,115%]". This corrected confidence interval is consistent with the listed p value of p<0.001 and the other statistics listed.

Color Key	
	Estimated risk/benefit percentage is significantly less than 100% (p-value less than or equal to 0.05 and the injury risks of the device are much less than the injury benefits of the device), and the estimated net benefit is significantly greater than 0% (i.e., significant benefit exists). May be a reasonable device assuming acceptable functional characteristics.
	Estimated risk/benefit percentage significantly more than 100%) p-value less than or equal to 0.05 and the injury risks of the device are greater than injury benefits of the device), and the estimated net benefit significantly less than 0% (i.e., significant disbenefit exists). Device causes significantly more injuries than it prevents.
	Estimated risk/benefit percentage not significantly different from 100% (p-value is greater than 0.05, the injury risks of the device are effectively equal to injury benefits of the device), and the estimated net benefit percentage not significantly different from 0% (within 95% confidence interval). Injury risks of the device are effectively equal to injury benefits of the device.

In Table 1, it can be seen that in comparison to the unhelmeted baseline, the probability of fatality outcomes for unhelmeted with CPD are very similar to those for normalized injury cost. The results indicate that the Quadbar has no statistically significant net benefit, and in fact the uncertainty (i.e., the confidence interval) is larger for the probability of fatality than for the normalized injury cost. Likewise, p-value (0.202) is larger for the unhelmeted/CPD probability of fatality results, indicating that, in terms of the probability of fatality, the probability that the Quadbar has a net harmful effect on the population of overturns is 0.10 (i.e., a one in ten chance).

Likewise, for the full face helmet and the half helmet, inclusion of the probability of fatality data in Zellner, et al. (2014a) and Zellner, et al. (2014b) did not change the conclusions regarding the effects of the Quadbar, which is that it has a statistically insignificant effect on the population of overturns.

51. ***"As this review did not have access to the simulation videos, it is not possible to comment in detail on the particular scenarios or initial conditions which led to these or any other simulated fatalities." (p23, paragraph 2)***

See comments in Section V.12.

SECTION VI

COMMENTS ON "CONCLUSIONS" (SECTION 5)

The comments on the Wordley (2012) "Conclusions" (Section 5) are as follows.

1. ***"Unfortunately, the video outputs from the new simulations were not made available for review, hence it is not possible to comment on the effects that these changes may have had on the scenarios modelled, or the level of realism attained." (p24, 1st bullet)***

   As discussed in Sections V.12 and 13, this statement could easily give rise to misunderstanding as suggesting that DRI has somehow resisted examination of its research. In fact, unlike all other researchers in this field, in the previous DRI research, DRI made available (during the course of the 2002- 2009 Victorian inquests) *all* of its several hundred simulation input files, *all* of its output files including videos of the several hundred simulations that had been done by that time, and its proprietary simulation software.

   Subsequently, and most recently, thousands (i.e., 3,080) of updated simulations have been run, and video conversions of the corresponding animations do not exist at this time.

   Following internet publication of the DRI updated (i.e., Zellner, et al. (2012)) research report, Wordley's ISCRR institute inquired as to the availability of simulation "videos". DRI sent the following reply in October 2012, which is still valid and applicable:

   > "At the current time, there are no "videos" of the simulations. Computer-generated visual displays (i.e., "animations") of the 3,080 simulations can be generated from the output files, and they have been reviewed in detail, but they require licensed software to view. Creating videos would involve a significant effort, and extensive (i.e., perhaps terabytes) of storage capacity.
   >
   > More importantly, currently, we are not aware of any suitable scientific forum or mechanism for scientific exchange in this area. If and when such a forum and mechanism exists, we are willing to provide access to the simulations, to appropriately qualified researchers."

   In short, making videos of the simulations "available" is not a small thing, and would require hundreds of DVDs and an extensive effort to generate them. DRI is willing to make the underlying "visual displays" available to qualified researchers for inspection, if/when there is a suitable scientific forum. Likewise, for such a forum, DRI is willing to provide example videos drawn from the 3,080 simulations.

In addition, understanding what is visible in the simulation "visual displays" is highly dependent on the viewing angle, and an understanding of what information is in the associated (but not explicitly modeled) UK/US case. Moreover, as extensively discussed in Zellner, et al. (2012b), previous comments and efforts to evaluate the "level of realism" based on viewing videos from the previous research were very often inaccurate, misinterpreted, without clear basis, irrelevant, or based on false premises. As noted in Section V.12 above, however, in principle DRI is willing to provide access to the animations and/or videos.

2. ***"A new rider asphyxiation criterion was developed and applied to all simulations. Only a small percentage of asphyxiation injuries were predicted, and the same likelihood of occurrence was observed both with and without the CPD added. In the CPD equipped cases, such incidents were almost exclusively due to the Quad bike coming to rest on its side, on top of the rider. Further investigation of the video results outputs should be conducted to determine if this injury mode is in fact realistic, or an artefact of other changes made to the modelling, such as the increased hand grip force." (p 24, 2nd bullet)***

   In regard to the statement that "only a small percentage of asphyxiation injuries were predicted", as noted in Section VI.2 above, there are no videos, although the animations could be used for this purpose. This statement is unclear in regard to exactly how the visual results could clarify whether "this injury mode is…an artefact of other changes made to the modeling."

3. ***"The increased sensitivity of the unhelmeted rider to head injuries, compared with other regions of the body, may have contributed to a relative over prediction of the safety benefits due to the use of helmets, and the relative under estimation of safety benefits due to other devices, including the CPD" (p 24, 3rd bullet)***

   This statement is speculative and without basis. One would certainly expect that there would be "increased sensitivity of the unhelmeted rider to head injuries, compared with other regions of the body", specifically due to absence of a helmet liner which can absorb energy (as related to head accelerations and brain injury potential) and lack of coverage of the skull (as related to force distribution and skull fractures). However, the manner in which this increased sensitivity might have "contributed to relative over-prediction of the safety benefits due to the use of helmets" and the relative under estimation of safety benefits due to other devices, is speculative, without basis, and unclear.

   This statement is also out-of-date, because Zellner, et al. (2014b) focused on the effects of a full-face helmet and a half helmet (in addition to the unhelmeted condition) on CPD effectiveness, as well as helmet effectiveness. As discussed in Section V.18, in terms of normalized injury costs, the net benefit of the half helmet (in comparison to no helmet) with a Baseline ATV was found to be a statistically significant 49%

(36%, 62% confidence interval), whilst for the full face helmet it was a statistically significant 60% (45%, 75% confidence interval). These helmet effectiveness estimates are in the range of helmet net benefits based on accident data (for all types of helmets), for example as reported by, for example, Rodgers (1990) of the CPSC, who estimated that helmet wearing on ATVs reduces the risk of non-fatal head injuries by 64%. Therefore, the Wordley (2012) statement that inclusion of skull fracture for uncovered portions of the head resulted in "over-prediction of the safety benefits…of helmets" is unsupported, and in fact refuted by these results.

Likewise, the Zellner, et al. (2014b) results (see Table 1 above) indicate that neither the presence of a helmet, or the style of helmet, have a statistically significant effect on CPD effectiveness. The Quadbar was found to have no statistically significant net benefit, regardless of the extent of helmet coverage. The Quadbar net benefit was found to be:

- A statistically insignificant 12% (-5%, 29%) (p=0.153) for unhelmeted;
- A statistically insignificant -20% (-42%, 2%) (p=0.066) for the half helmet;
- A statistically insignificant -3% (-24%, 18%) (p=0.760) for the full face helmet;

Therefore, it can be said that the degree of helmet coverage, and the possibility of the uncovered skull being fractured, had no discernible effect on Quadbar effectiveness, which was found to be statistically insignificant for all helmet /no helmet configurations.

Note that in these results from Zellner, et al. (2014b), the configuration closest to having a statistically significant outcome was the half helmet (p=0.066) (p=0.05 being the commonly recognized level for being "somewhat convinced" (Box, et al (1978)) of the sign of the outcome), and in this case the outcome was in the direction of the Quadbar having a negative net benefit, i.e., a net harm.

4. ***"A significant difference was noted between the style/standard of helmet being recommended by the industry and government bodies for Quad bike use (i.e. compliant with AS/NZS:1698), and that which was simulated by DRI. The current Australian standard allows a relatively light weight, half helmet style, whereas DRI simulates a full-face style helmet more commonly used by on-road motorcycle riders. Therefore the estimates of increased safety due to the use of a helmet, as quoted by DRI, should be considered as a very best case scenario. It is recommended that the modelling and simulation of an AS/NZS:1698 compliant helmet be utilised for future Quad bike accident simulations." (p24, 4th bullet);***

As discussed in Sections II.6, V.16 - 19, V.49 and VI.4, this statement is inaccurate and represents a major misunderstanding in Wordley (2012).

First, the full-face Bieffe B12 helmet is fully compliant with the impact and coverage requirements of AS/NZS:1698, as well as with those of ECE R22.[64] Second, AS/NZS:1698-2006 does not prescribe any particular style/type of helmet; both full-face and open-face motorcycle helmets that are universally used by motorcyclists throughout Australia comply with AS/NZS:1698, as do some half-helmets (also known as "shorty" helmets) such as the THH T70 suggested by Wordley (2012). Third, the ATV manufacturers' recommendations regarding helmets are that they comply with the applicable federal, state and/or other standards[65] and are consistent with the guidelines provided by the ATV Safety Institute as follows:

> "Select a helmet that meets or exceeds your state's safety standards and carries either the Department of Transportation (DOT) label or the Snell Memorial Foundation label. Your helmet should fit snugly and fasten securely. Full-face helmets help protect your face as well as your head. Open-face helmets are lighter and may be cooler, but should be used with mouth protection. Eye protection should be used with both types of helmets."[66]

Note that the only ATV helmets illustrated in the aforementioned guideline are full-face helmets. Also note that inquiries indicate that neither the FCAI nor the largest Australian ATV importer specifically recommend "shorty" (i.e., half) helmets.

This statement is also out-of-date since, as further discussed in Section VI.4 and in Zellner, et al. (2014b). Substituting the half-helmet THH T70 characteristics into the DRI simulation and re-running the simulations and risk-benefit analysis indicates that the Quadbar risk/benefit and net benefit outcomes are <u>similar</u> to those with the full-face (Bieffe B12) helmet, and do not change any of the conclusions regarding the Quadbar. As might be expected, the net benefit of the half-helmet alone (as a protective device) is less than that of the full-face helmet alone, but still substantial. Note that Rodgers (1990) of the CPSC reported net benefit of 64% for ATV helmets for non-fatal head injuries, which are similar to the net benefits predicted for full-face helmets by the DRI simulations (i.e., 60% for predominantly non-fatal injuries).

5. ***"These changes [in injury AIS coding] eliminated approximately 25 reported "trunk" injuries and 15 "arm" injuries. Approximately one third of the final 49 injuries retained were head injuries. The overall validity and usefulness of a correlation resulting from such a heavily filtered dataset was considered minor.***

---

[64]See Footnote 2.
[65]See Footnote 3.
[66]Anon., "Tips and Practice Guide for the All-Terrain Vehicle Rider", page 4, www.atvsafety.org/InfoSheets/ATV_Riding_Tips.pdf, ATV Safety Institute, a Division of the Specialty Vehicle Institute of America, (accessed 9 July 2013).

The potential for researchers [to] tune their AIS coding of injuries to match their simulated results was also noted." (p24, 6th bullet);

As discussed in Sections IV.6 and V.21, this statement could easily give rise to misunderstanding as somehow involving an inappropriate procedure.

The purpose of the injury coding in the subject research was *to enable calibration of the simulation mode*, which, at the current state of technology, is able to predict only a limited set of locations, types and severities of injury (as is the case for all crash dummies applicable to cars, motorcycles, etc., worldwide). It would therefore be *inappropriate, senseless and not possible to attempt to include those injuries that cannot be monitored by a crash dummy*. To include those injuries that cannot be monitored would be to attempt to compare "apples to oranges". This is the appropriate reason that some locations, types and severities of injuries were "filtered" (i.e., removed) from consideration.

As further discussed in Section V.30, in regard to "the potential for researchers to tune their AIS codings", this statement is speculative, defamatory, inaccurate and without any basis. Such a process would have been highly inappropriate (not to mention extraordinarily difficult). The codings were simply updated to remove invalid (i.e., non-ATV) cases; to remove body regions that cannot be monitored with existing dummy technology; and to update the codings to the 2005 AIS coding conventions. The accusation implied by this statement is refuted.

6. ***"Predicted Injury/Benefit ratios from this updated work have been presented using a "single baseline" method rather than the commonly accepted and more logical "multiple baseline" method."(p25, 1st bullet)***

As discussed in Section II.8 and elsewhere in this report, this statement is inaccurate, as both the "single baseline" results and the "two baseline" results were presented in Section IV.B and in Appendix I, respectively, of Zellner, et al. (2012). This is also true of the later published papers, i.e., Zellner et al. (2014a) (in which single baseline results are presented in Appendix B and two baseline results are presented in Appendix C) and Zellner, et al. (2014b) (in which both single baseline and two baseline results are presented in the main text). The main text discussion of Zellner, et al. (2012), which is based on the "single baseline" method, cross-references (on page 64) the Appendix I "two baseline" discussion.

In addition, this statement is inaccurate as the "two baseline" method is *not* more "commonly accepted" compared to the "single baseline" method: there are no standards requiring it, and there is no other original research in this field other than DRI's that uses a "two baseline" method. The single baseline method is relevant because (a) it is relative to the intended (safer) condition (i.e., helmeted) and (b) it can capture what occurs if a user were to falsely assume that the CPD has a significant protective effect and that he/she no longer needs to wear a helmet. As further discussed in Section V.36 herein, the "two-baseline" approach can be misleading in that it shifts focus away from the main hazard, which in this case is not wearing a helmet. Not wearing a helmet is approximately 2.5 times more hazardous than wearing a helmet, and even if a device such as a CPD reduced that hazard by a small but statistically significant amount – which the DRI data do not indicate – the unhelmeted hazard – and the large benefits of wearing a helmet – far exceed the estimated effects of fitting a CPD while unhelmeted, which were found by the DRI research to be statistically insignificant.

In addition, the validity of the statement that the two baseline method is "more logical" depends on the reader's perspective. If the reader is an ATV safety advocate or ATV manufacturer, who is recommending that all riders wear helmets, then a single baseline, the intended "helmeted" baseline ATV condition, is more logical.

7. ***"Recalculation of the injury risk/benefit results using the more correct multiple (or matched helmet condition) baseline method shows that, according to updated DRI's research, the use of a CPD provides a risk/benefit ratio of 108% [69%, 169%] (marginally less safe) for the helmeted condition, and 68% [42%, 114%] (slightly safer) for the unhelmeted condition."(p 25, 2nd bullet)***

As discussed in Section II.10 and elsewhere in this report, this statement is inaccurate, incomplete and somewhat misleading because the terms "slightly safer" and "marginally less safe" [or, more accurately, "marginally more dangerous"] (and the 68% and 108% values) *only* apply to the DRI *simulation sample*. This sentence should say "slightly safer for the *simulation sample*" and "marginally less safe [or marginally more dangerous] for the *simulation sample*". As clearly stated in the DRI report, the results for *all samples*, i.e., the "population estimate", are that the effects of the Quadbar are statistically insignificant, *for both the helmeted and the unhelmeted condition*, i.e., using the "two baseline" method that Wordley (2012) seems to prefer.

Wordley (2012) cannot have it both ways: Wordley (2012) cannot simultaneously *reject* DRI's "simulation sample" as being invalid, which is stated in the Wordley (2012) report,[67] and at the same time *accept the specific "sample mean" results* that apply *only* to DRI's simulation sample, and not to the population of overturns.

As noted, the overall population results (as opposed to the specific simulation sample results) reported by DRI are that for the *population estimate* applicable to *all* samples of ATV overturns, the effects of the Quadbar are *statistically insignificant*, within a 95% confidence interval *for both the helmeted and the unhelmeted condition*. A 95% confidence interval (equivalent to p=0.05, Box, et al. (1978)) is by far the most widely accepted standard for being "somewhat convinced" of the reality of an outcome.

[67]See Footnote 11.

8. ***"Due to the fact that the 95% confidence intervals (listed in the square brackets) for both these results were observed to straddle the neutral risk/benefit value (100%), both findings were considered statistically insignificant by the DRI researchers. It was noted though, that for the unhelmeted condition the CPD was very close (within 14%) to returning a statistically significant prediction of increased safety." (p25, 2nd bullet);***

   As discussed in Section V.45, the statement that "the addition of the Quadbar is very close (within 14%) to returning a statistically significant prediction of increased safety is inaccurate and potentially misleading. The subsequent Zellner, et al. (2014b) Table 1 results, corresponding to the Zellner, et al. (2012a) report being commented on, indicate that the associated p-value for the unhelmeted case is p=0.153, which according to Box, et al. (1978), is three times "conventional "critical" significance level in common use" of p=0.05 where one can be "somewhat convinced" that the Quadbar has any net effect, based on these results..

   In addition, the statement that "the confidence limits for the helmeted case straddle the neutral value" is unclear in its meaning and implication and could be easily misunderstood as implying that somehow the results were "neutral". The "helmeted case" (i.e., the Quadbar/helmeted in comparison to the Baseline ATV/helmeted, also described in Zellner, et al. (2014b) Table 1) indicates that (for the normalized probable injury cost) p=0.760. This probability level is more than 15 times the commonly accepted probability level at which one is "somewhat convinced" that the Quadbar will has some non-zero net effect (i.e., net benefit or net harm), based on these results. Furthermore, if the device does have some non-zero net effect, this result can also be interpreted as indicating that there is a 0.620 (=1-p/2) probability that the Quadbar is in fact harmful for the population of overturns, which far exceeds an acceptable probability of harm. It is not reasonable, prudent or ethical to recommend a device with this much potential to worsen the injury outcome.

9. ***"Despite the many changes made to the simulation methods used in the updated work, the true updated results are in fact very similar to the initial study." (p 25, 3rd bullet)***;

   This statement fails to observe and acknowledge that the "many changes made to the simulations methods" giving "very similar results" is an indication of the stability of the analyses, i.e., not being overly sensitive to changes in methods and assumptions.

   This statement also fails to observe and acknowledge that the "many changes made to the simulation methods", which were in part made to address the criticisms of the previous work which Wordley (2012) describes as "serious limitations", in fact according to Wordley (2012) had little effect on the outcomes, i.e., the “serious limitations” were perhaps not actually as serious as they were stated to be by Wordley and Fields (2012).

SECTION VII

COMMENTS ON "RECOMMENDATIONS" (SECTION 6)

The comments on the Wordley (2012) "Recommendations" (Section 6) are as follows:

1. ***"Given that the current review has found no new evidence to contradict the previous conclusions arising from the ISCRR Quad bike Safety Devices Snapshot Review1, the prior recommendations can be ratified with only minor updates." (p 26, paragraph 1)***

   This statement totally ignores the fact that all of the so-called "serious limitations" and "critical limitations" (noted by Wordley (2012), p1 and p4 respectively) that were identified in the previous work were addressed in the updated research, either by being fully explained (e.g., as being due to misunderstandings by the critics), or if arguably valid, were fully accounted for in the Zellner, et al. (2012a) updated research, as discussed in Sections II.1, III.1 and VI.10.

   The fact that the updated analysis gave the same overall conclusions regarding the Quadbar, after having addressed all of these "serious limitations", can only be attributed to the "serious limitations", and their correction, not having as large an effect as thought by Wordley (2012). This fact should have been noted by Wordley (2012), and should have been the basis for him reconsidering the findings.

   The "prior recommendations" should have been substantially changed, as a recommendation for "the use of appropriately tested CPDs…for riders who use Quad bikes at low speeds on farms and in other work places" is without basis, having regard to the updated DRI research, which indicates, for example, that the Quadbar/helmeted configuration has a 0.620 probability of being harmful in comparison to the Baseline ATV/helmeted configuration for the population of overturns. This far exceeds an acceptable probability of harm. It is not reasonable, prudent or ethical to recommend a device with this much potential to worsen the injury outcome.

2. ***"A new incident dataset be developed based on Australian and perhaps New Zealand Quad bike fatality reports, and a range of simple generic roll scenarios (as described below). This dataset should be used for future simulations into the effectiveness of crush protection devices." (p 26, item 2)***

   This statement is incomplete as it does not explain why the scenarios should be "simple". If the injurious phenomenon is found to be inherently complex, then it stands to reason that a test of countermeasure effectiveness should be as complex as is necessary to capture the phenomenon of interest. Examples of this are the US/NCAP passenger car Rollover Resistance tests (which is complex, requiring a steering robot to implement) and the US FMVSS 214 test of passenger car yaw stability and steer-ability, also requiring a steering robot to implement.

The fact is, Wordley (2012) and others have not yet constructed the required "incident dataset" in sufficient detail (or in any detail) that would be needed in order to understand how ATV overturn injuries and fatalities actually occur, in terms of speeds, slopes, directions of overturn, heading angles relative to slopes, and so on; and therefore are not in a position to recommend whether they need to be "simple" or "complex".

This statement recommends the use of "generic" overturn scenarios, which is what DRI used in its simulation analyses, and in its 1998 full-scale overturn tests (Van Auken, 1998) with instrumented crash dummy; for which DRI has been criticized. This statement seems inconsistent in regard to what has been previously stated by Wordley (2012) and others who have criticized the DRI work. If use of "generic" overturn scenarios is now recommended, on what basis is it recommended, and is that not a "new finding" that would "contradict the previous conclusions" of Wordley and Field (2012), as discussed in Section VI.1?

3. ***"A preliminary standard be proposed for the design and specification of Quad bike CPDs, perhaps initially based upon those for Tractor ROPS, or the performance of the CPD when subjected to the experiments proposed below." (p 26, item 3) (p 27, paragraph 1)***

This statement is without basis, as it incorrectly pre-supposes that:

- some scientifically valid or comprehensive data currently exist that indicate that a CPD on an ATV has more injury benefit than injury risk, across a realistic range of overturn conditions, when in fact there are none;
- tractor ROPS are in some respect comparable with or relevant to the design of a CPD for an ATV, whereas:
  - (a) under Australian state law, ROPS are not required for tractors of less than 560 kg mass[68], while a typical ATV has a mass of only 275 kg; and
  - (b) tractors are totally different types of vehicles from ATVs, as tractors with ROPS are not "rider active" vehicles, have much greater size and mass, may have large unguarded wheels in close proximity to the driver, typically have seat belts and have an entirely different type of occupant seat, and have very different rollover motions and occupant motions.

[68]For example, Chapter 5, Part 5.1, Division 7, Subdivision 3, Clause 216 of the current NSW Work Health and Safety Regulation 2011. Also Australian Standard AS 1631.1-1996.

4. ***"Experiments: Basic Quad bike lateral roll, forward flip and back flip tests for a range of speeds and slope angles should be conducted using an instrumented dummy. The 'Vehicle Accelerator' developed by The University of Southern Queensland could be utilised for these tests. Such tests should confirm and quantify the level of injury protection provided by CPDs for these incident types." (p 26, item 5, paragraph 2))***

   This statement is without basis and inappropriate. Although "lateral roll, forward flip and backward flip tests for a range of speeds and slope angles" might be reasonable, provided they are of sufficient range and variety, the USQ "Vehicle Accelerator" is not a suitable device for implementing these tests. As discussed in Zellner, et al. (2011b), the Snook test methodology suffered from a number of significant deficiencies including that:

   - It provides unrealistic ATV motions (i.e., real ATVs generally do not accelerate sideways, slide sideways at relatively high speed, on a frictionless surface, and then "trip" on a mechanical stop;

   - It has no documentation which supports the claim that it provides "linear and angular momentum *commensurate* with those experienced in an actual roll over event"[emphasis added];

   - The unrealistic *sliding sideways at relatively high speed on a frictionless surface*, and then "tripping" on a mechanical stop means that the initial linear momentum (and therefore the angular momentum) provided by the apparatus for the sideward test by definition cannot be "typical", "correct" or "commensurate";

   - It produces an angular momentum induced by the lateral "trip" mechanical stop that would be far greater than those that have been observed in that real ATV low speed overturn tests, described by, e.g., Van Auken, et al. (1998), which tend to be far more gradual;

   - it provides a sudden, impulsive, artificial, relatively high speed, "tripped" rotation;

   - it has not been demonstrated for forward pitch overturns, or for combined axis overturns, such as yaw/roll, pitch/roll, yaw/pitch/roll, which occur in real ATV overturns, and which are relevant;

   - it is based on a "stop" bar, which impacts against the tyre sidewalls and/or tread, in order to "trip" the ATV. Therefore the angular speed resulting from the trip strongly depends on the particular ATV's tyre characteristics (e.g., stiffness, damping, inflation pressure, etc.). This means that one could change the outcome of the test of a CPD merely by changing the tyres used, which is not reasonable;

– Uses a surrogate handlebar, which is *not* part of the original ATV, which appears to be able to arrest the roll, and which has been observed to hold the Quadbar off the ground.

DRI has suggested, and has conducted in the past (i.e., Van Auken (1998)), overturn tests with a remotely-controlled ATV (enabling repeatable steering, braking and/or throttle inputs) on natural terrain, with an instrumented injury monitoring dummy compatible with straddle seat, handlebar, helmet-required vehicles, able to monitor for injuries in several directions and across many body regions.[69] Efforts are needed to demonstrate and to maximize the repeatability and reproducibility of this methodology.

5. ***"Simulations: Computer simulations should be used to accurately correlate the experimental tests, on the basis of both dummy and bike motions, and recorded injuries. Providing adequate correlation can be achieved, these simulations can be extended to incorporate additional generic overturn events (which are more difficult to reproduce experimentally) and the proposed new injury dataset containing Australian fatality scenarios." (p26, item 5, paragraph 3)***

   Effectively, this is exactly what DRI has done, for which it has been criticized, by Wordley and Field (2012), Wordley (2012) and others. If simulations that accurately correlate with experimental test are now recommended, on what basis are they recommended, and is that not a ***"new finding"*** that would ***"contradict the previous conclusions"*** of Wordley and Field (2012), as discussed in Section VI.1?

6. ***"Fit a large sample (>100) of Quad bikes currently being used by farmers with light weight and inexpensive devices which are capable of recording video, accelerations, speed and map position of the bike whilst it is being used in the field." (p26, item 5, paragraph 4)***

   While this is a good idea, in order to find any statistically valid outcomes in terms of overturns, it seems likely that the sample of ATVs would have to be much larger than n=100 vehicles, and/or involve data collections times exceeding one year, in order to capture sufficient number of overturns, as overturns are relatively rare events and injurious overturns are even rarer (e.g., as found by Van Ee (2012)). In addition, if the purpose is to capture "normal" (non-overturn) riding conditions, the results would be expected to be unrelated to riding conditions that result in an overturn, or that result in an overturn/injury or overturn/fatality.

7. ***"In the mean time [sic], regulatory bodies should consider recommending the use of appropriately tested crush protection devices (CPDs) for riders who use Quad bikes at low speeds in the workplace and on farms. Such devices have been shown to reduce the severity of Quad bike roll-over, and have the potential to reduce the***

[69]The ISO 13232 (2005) MATD dummy is currently the standardized crash dummy that has these capabilities.

injuries and fatalities associated with these loss of control events." (27, paragraph 1)

First, as discussed in Section V.49, if one accepts the specific DRI simulation sample (i.e., the "sample estimate") as representing all ATV overturns that occur, then it is not reasonable, prudent or ethical to recommend fitting a device (i.e., the Quadbar) that is predicted to worsen the overall injury outcome for those who wear helmets (i.e., for whom the injury risks of the Quadbar are greater than the injury benefits, by a ratio of 108%).

Second, as also discussed in Section V.49, if one instead accepts the validity of the "population estimate" based on the outcomes from the DRI simulation sample, then the results of the Zellner et al. (2012) research is that the Quadbar has no statistically significant net injury benefit. By way of comparison, if a medical device, medical procedure or pharmaceutical is found (in tests with human surrogates) to have *no statistically significant net benefit*, common practice is *not to proceed* with clinical trials on humans, or with widespread use of the device, procedure or pharmaceutical. In essence, despite these findings, Wordley (2012) is recommending that regulatory bodies proceed with such trials with humans, and with widespread use of the device.

Third, the Zellner, et al. (2012a) results indicate that, with a Quadbar, the likelihood of an ATV coming to rest on its side on a rider will be substantially *increased*, and it is known that ATVs coming to rest on their side are observed in Australian and New Zealand fatal and serious injury ATV accidents. See Footnote 2.

Fourth, the definition of "low speed" is unknown, no clear evidence of accident speed is available from ATV incidents and there is no known evidence that farmers only ride at "low speed", or can be forced to do so. Moreover there is no known evidence that CPDs such as a Quadbar are beneficial and not harmful at "low speeds", as opposed to medium speeds or high speeds. This recommendation appears to be without basis and based on conjecture.

SECTION VIII

SUMMARY AND CONCLUSIONS

This report provides comments on a report by Wordley (2012) entitled "Quad bike Crush Protection Devices (CPDs): Updates to ISCRR Snapshot Review C-I-12-022". This Wordley (2012) report supplements the earlier Wordley and Field (2012) report and is primarily a commentary upon two additional publications by Dynamic Research, Inc. (DRI) relating to the effectiveness of an Australian designed and manufactured CPD, the Quadbar™, which is proposed to be fitted to All Terrain Vehicles (ATVs). Those DRI publications are Zellner et al. (2012), "Updated Injury Risk/Benefit Analysis of Quadbar Crush Protection Device (CPD) for All-Terrain Vehicles (ATVs)" and Zellner, et al. (2012a), "Replies to Lower (2011) Comments".

Regrettably, most of the commentary in the Wordley (2012) report comprises interpretations and analyses of, in particular, the Zellner, et al. (2012) report, which are frequently inaccurate, somewhat misleading, based on false premises or misunderstandings of relevant technical issues, without basis or incomplete. In addition, many of the comments in the Wordley (2012) report are now out of date as a result of more recent DRI research (i.e., Zellner, et al. (2014a) and Zellner, et al. (2014b)).

The principal deficiencies in the Wordley (2012) report include that:

– Wordley (2012) bases many of its conclusions about the Quadbar entirely upon the invalid test methods of Snook (2009), in which an ATV was tested alone, without any rider surrogate. Consequently, those tests took no account of where a rider would be or how, and at what point, he would separate from the ATV in the event of an overturn but, rather, made baseless assumptions as to where a hypothetical rider "would be", with inexplicably different assumptions as for the baseline ATV and the Quadbar ATV vehicle configurations, including that a rider would *always* come to rest under an upside down (non-Quadbar) ATV.

  Snook (2009) also failed to make any direct measurement of injury assessment variables. Instead, Snook (2009) assumed, incorrectly, arbitrarily and without any indicated basis, that some particular outcomes would be injurious, or more injurious than other outcomes. For example, it was assumed that an ATV coming to rest upside-down was injurious, whereas an ATV coming to rest on its side was mistakenly[70] assumed to be harmless, and that an ATV continuing to roll "farther"

[70] For example, "A Handbook for Workplaces: Quad Bikes on Farms" (Work Safe Victoria, Edition No. 2, August 2009, p6) cites a case of "A 75-year-old man was killed whilst operating a quad bike equipped with a 50-litre spray tank full of chemical spray. He was working in a wet area on an incline of 20–30 degrees. The man's wife discovered him lying face down towards the rear of the quad bike which was on its side on top of the man." Another example was a New Zealand ATV on-its-side case described at the Technical Engineering Group meetings, Sydney, 5 to 7 October 2010, resulting in long term rider entrapment and subsequent medical amputation of the leg. The Peter Crole ATV fatal accident in Australia involved an ATV coming to rest on its side on the rider, with heavy equipment on front and rear cargo racks

(i.e., farther away from a rider who as a result of natural forces typically separates from the ATV at an early stage of an overturn) was more injurious than an ATV coming to rest in closer proximity to the rider, which are highly questionable assumptions for which there is contradictory evidence.

Despite the many serious flaws in the Snook (2009) research, Wordley (2012) states "***The limited number of experimental and simulation results which were considered valid [and primary among them, Snook (2009)] indicated that CPDs demonstrated the potential to reduce rider harm in low speed roll-over events." (p1, paragraph 2).***

– Wordley (2012) wrongly asserts that there is "***A significant difference between the style/standard of helmet being simulated (full-face) and that which is currently recommended by the industry and government bodies for Quad bike use (i.e. compliant with AS/NZS:1698)" (p1, 3rd bullet).***

  In fact:
  - the full-face (Bieffe B12) helmet simulated in Zellner et al. (2012) is *fully compliant* with the impact and coverage requirements of AS/NZS:1698-2006 (as well as ECE R22.[71]);
  - the ATV manufacturers' recommendations regarding helmets *do not prescribe a particular style/type of helmet* but, rather, that they comply with applicable federal, state and/or other standards[72] and that they are consistent with the guidelines provided by the ATV Safety Institute;

  In any event, the THH T70 half helmet suggested by Wordley (2012) was used by Zellner et al (2014b) in their extended simulation analysis, and this resulted in no change to the overall conclusions about the Quadbar in comparison to the Zellner et al. (2012) research.

– Contrary to the statement in Wordley (2012) that "***Predicted Injury/Benefit ratios from this updated work have been presented using a "single baseline" method rather than the commonly accepted and more logical "multiple baseline" method." (p1, 3rd bullet),*** in fact both the "single baseline" results and the "two

(Rechnitzer and Grzebieta, 2004). In Zellner et al. (2004), Appendix E, two of the three UK/US asphyxia cases involved unknown resting orientations, and the third involved wheel/face contact (which cannot be an upside-down orientation). In his findings in the inquest into the death of H, a child, Queensland Deputy State Coroner Lock was unable to ascertain whether an ATV had rolled ¼ times or 1¼ times onto its side (although he found it "difficult to conceive the latter figure"), but either way, the child had died of asphyxia after being pinned beneath the side of the ATV. In addition, Figures 10 and 11 in Zellner et al. (2012) indicate that based on 3,080 overturn simulations, the baseline ATV came to rest more frequently on its side, and that adding a Quadbar more than doubled the number of pinned/on-its-side/predicted asphyxiations.

[71] Smith, T.A, Koerner, J.E., Certification Test Report according to AS NZS:1698-2006, DRI-TM-14-67, 16 September 2014.

[72] For example, in the US, US/DOT/FMVSS 218 (which is very similar to AS/NZS:1698) which likewise does not specify the style of helmet.

baseline" results were presented in Zellner, et al. (2012), in Section IV.B and Appendix I, respectively. This comment is also now out of date as both the "single baseline" results and the "two baseline" results were presented *within the same portion of the text*, in the most recent Zellner, et al. (2014b) published paper. In addition, this statement is inaccurate as the "two baseline" method is not more "commonly accepted" compared to the "single baseline" method: there are no standards requiring it, and, far from being "commonly accepted", there is in fact no other original research in this field, other than DRI's, that uses a "two baseline" method.

– Despite having previously (in Wordley and Field (2012)) rejected the *specific simulation sample* used by DRI, Wordley (2012) relies upon that sample to inaccurately and misleadingly claim that "***according to DRI's updated research, the use of a CPD actually returns a risk/benefit ratio of 68% [42%, 114%] (slightly safer) for the unhelmeted condition, and 108% [69%, 169%] (marginally less safe) for the helmeted condition." (p2, paragraph 2)***.

  This claim ignores the fact that the results for the *population* of overturns reported in Zellner et al. (2012) indicate that the effects of the Quadbar are, *for both the helmeted and the unhelmeted condition, statistically insignificant* within a 95% confidence interval, the latter confidence interval (equivalent to p=0.05, Box, et al. (1978)) being by far the most widely accepted standard for being "somewhat convinced" of the reality of an outcome. The fact that the square bracketed percentages span the 100% value indicates that the 68% and 108% ratios (the point estimates for the simulation sample only) are, for the population of all overturns, *statistically indiscernible* from 100%, i.e., that they are in fact *statistically insignificant*. They are categorically *not* "slightly safer" or "marginally less safe", which descriptions would *only* apply to the specific UK/US simulation sample that Wordley and Field (2012) rejected as being invalid.

– The statement by Wordley (2012) that ***"for the unhelmeted condition the CPD was very close (within 14%) to returning a statistically significant prediction of increased safety."(p 25, 2nd bullet)*** is contrary to standard practice in statistical analysis, which does not involve stating that a factor is "very close" to a statistically significant prediction. The statistical test in and of itself returns either a statistically significant or a statistically not significant outcome. The power of the statistical test confirms the reliability of the outcome, which in this case was a non-statistically significant difference between the risk/benefit ratio with a CPD and without a CPD, for both helmeted and unhelmeted conditions.

  Moreover, the fact that the probability value (p=0.153) for the Quadbar unhelmeted risk/benefit outcome is *more than three times* the "conventional 'critical' significance level in common use" (i.e., p=0.05, Box, et al (1978)) indicates that the Quadbar was *not* "very close" to returning a statistically significant prediction of increased safety.

– The suggestion by Wordley (2012) that ***"A preliminary standard be proposed for the design and specification of Quad bike CPDs, perhaps initially based upon those for Tractor ROPS, or the performance of the CPD when subjected to the experiments proposed below." (p 26, paragraph 1)*** incorrectly pre-supposes that:
    - some scientifically valid or comprehensive data currently exist that indicate that a CPD on an ATV has more injury benefit than injury risk, across a realistic range of overturn conditions, when in fact there are none;
    - tractor ROPS are in some respect comparable with or relevant to the design of a CPD for an ATV, whereas:
        (a) under Australian state law, ROPS are not required for tractors of less than 560 kg mass[73], while a typical ATV has a mass of only 275 kg;; and
        (b) tractors are totally different types of vehicles from ATVs, as tractors with ROPS are not "rider active" vehicles, have much greater size and mass, may have large unguarded wheels in close proximity to the driver, typically have seat belts and have an entirely different type of occupant seat, and have very different rollover motions and occupant motions.

– In stating that ***"The overall validity and usefulness of a correlation resulting from such a heavily filtered [accident] dataset was considered minor. The potential for researchers [to] tune their AIS coding of injuries to match their simulated results was also noted" (p27, paragraph 5)***, Wordley (2012) misconceives the purpose of the injury coding in the subject research, which was *to enable calibration of the simulation model.* This filtering is necessary and entirely appropriate because, at the current state of technology, simulation models are able to predict only a limited set of locations, types and severities of injury (this is the case for all crash dummies worldwide, including those applicable to cars, motorcycles, etc.). It would therefore be *inappropriate, senseless and impossible to attempt to include and calibrate those injuries that cannot be monitored by a crash dummy, which is what Wordley (2012) appears to recommend doing*. To include injuries that cannot be experimentally monitored and verified would be to attempt to compare "apples to oranges". This is the appropriate, and only, reason that some locations, types and severities of injuries were "filtered" (i.e., removed) from consideration.

  It is to be regretted that the failure of Wordley (2012) to appreciate the previous point has apparently prompted the accusation of impropriety implied in the statement, ***"the potential for researchers to tune their AIS codings…was noted***". This accusation is baseless and defamatory and is refuted.

[73]For example, Chapter 5, Part 5.1, Division 7, Subdivision 3, Clause 216 of the current NSW Work Health and Safety Regulation 2011. Also Australian Standard AS 1631.1-1996.

- The recommendation by Wordley (2012) that ***"In the mean time [sic], regulatory bodies should consider recommending the use of appropriately tested crush protection devices (CPDs) for riders who use Quad bikes at low speeds in the workplace and on farms." (p27)***, is without basis and inappropriate because:
    - the highly deficient Snook (2009) tests did not indicate any ***"potential to reduce harm"*** in any scientifically valid way;
    - the updated Zellner et al. (2012) research and overall findings, which addressed all of the supposed ***"serious limitations"*** previously suggested by Wordley and Field (2012), found that the Quadbar CPD has *statistically insignificant effects* on injuries and fatalities (as well as a risk/benefit percentage that is larger than that for any known vehicle safety device); and
    - there are no scientifically valid data which indicate that CPDs in general, or the Quadbar in particular, are more effective and less harmful at low speeds than at any other speeds.

Absent, therefore, any evidence to support the safety or efficacy of the Quadbar, the recommendation that they be fitted nonetheless is of doubtful ethical standing.

REFERENCES

Anon., The Abbreviated Injury Scale, Association for the Advancement of Automotive Medicine, 1990 Revision (1998 Update), 2005 Revision (2008 Update), Illinois, USA.

Anon., AS 1631.1-1996: Tractors – Roll-over protective structures – Criteria and tests, Part 1: Conventional Tractors, Standards Australia, Homebush, NSW, 1996.

Anon., ISO 13232: Motorcycles - Test and analysis procedures for research evaluation of rider crash protective devices fitted to motorcycles**,** 2005 Revision, Parts 1 – 8, International Organisation for Standardization, Geneva, 2005.

Anon., Motorcycles - Common Methodology for in-depth accident investigations, Parts 1 - 13, Organisation for Economic Cooperation and Development, Paris, 2002, revised 2007.

Anon., A Handbook for Workplaces: Quad Bikes on Farms, WorkSafe Victoria, Edition No. 2, Melbourne, August 2009.

Anon., NCIS Database Search, Australian All Terrain Vehicle / Quad bike Deaths, 1 July 2000 – 31 July 2009, compiled by National Coroners Information System, August 2009.

Anon., American National Standard for Four Wheel All-Terrain Vehicles, ANSI/SVIA 1 – 2010, Specialty Vehicle Institute of American, Irvine, California.

Anon., Report of the Technical Engineering Group (TEG) to the Heads of WorkplaceHealth and Safety Authorities (HWSA)**,** Trans Tasman Working Party on "Quad Bike Safety" Meeting, Sydney, October 5-6, 2010.

Anon., 2009 Public Release of the All-Terrain Vehicles Deaths (ATVD) Database, Consumer Product Safety Commission, Bethesda, Maryland, December 2010.

Anon., NSW Work Health and Safety Regulation 2011, New South Wales, (http://www.legislation.nsw.gov.au/inforcepdf/2011-674.pdf?id=8007be51-e708-4bbd-82e3-736a941fa8c6, accessed 2014-08-18)

Anon., American National Standard for Recreational Off-Highway Vehicles, ANSI/ROHVA 1-2011, Recreational Off-Highway Vehicle Association. (ROHVA), Costa Mesa, California, 2011.

Anon., Statistical Report: International road safety comparison - 2010, Department of Infrastructure and Transport, Bureau of Infrastructure, Transport and Regional Economics, Canberra, 2012.

Box, B.E.P., Hunter, W.G., Hunter, J.S., Statistics for Experiments, John Wiley & Sons, 20 New York, 1978.

Gravetter, F.J., and Wallnau, L.B., Statistics for the Behavioral Sciences, Fourth Edition, West Publishing, Minneapolis, MN, 1996.

Grzebieta, R. and Achilles, T., Report on the Quad-bar in Relation to ATV Rollover Crashworthiness, Department of Civil Engineering, Monash University, Australia, April 2007.

Iijima, S., et al. “Exploratory Study of an Airbag Concept for a Large Touring Motorcycle” Paper No. 98-S10-O-14, Sixteenth International Technical Conference on the Enhanced Safety of Vehicles, Windsor, Canada, 1-4 June 1998.

Kebschull, S.A., Zellner, J.W., Van Auken, R.M., and Rogers, N., “Injury Risk/Benefit Analysis of Motorcyclist Protective Devices Using Computer Simulation and ISO 13232”, Proceedings of the 16th International Technical Conference on the Enhanced Safety of Vehicles, Windsor, Canada, Paper Number 98-S10-W-26, 1-4 June 1998.

Lambert, J., Version 4 of a paper reviewing: Quad bike design, Computer simulation of Quad bike incidents as reported in various papers and presentations by Dynamic Research Inc., Quad manufacturers cynicism in respect of safety; and Quad bike safety improvements, John Lambert and Associates, Wandana Heights, Victoria, 2011.

Muñoz, S., Van Auken, R.M., Zellner, J.W., An assessment of the effects of the Robertson V-Bar ROPS on the risk of rider injury due to overturns resulting from ATV Misuse, Dynamic Research Inc. Technical Report DRI-TR-07-14, Torrance, California, 2007.

Parker, S.P. (Ed), McGraw-Hill Dictionary of Science and Engineering, McGraw-Hill, New York, 1984.

Rechnitzer, G. and Grzebieta R., Engineering Investigation regarding the death of Peter Crole, involving an All Terrain Vehicle on a property at Don Valley, on the 17th February 2002, DV Experts, January 2004.

Rechnitzer, G., Day, L., Grzebieta, R., Zou, R. and Richardson, S., All Terrain Vehicle Injuries and Deaths, Monash University Accident Research Centre, 19 March 2003.

Rodgers, G.B, “The Effectiveness of Helmets in Reducing All-Terrain Vehicle Injuries and Deaths,” Accident Analysis and Prevention, 1990 Feb; 22(1):47–58.

Smith, T.A., Satrom, P.R., Van Auken, R.M., Wiley, K.D., Test and Analysis Procedures for Measuring Skull Fracture Injury Probability, Dynamic Research, Inc. Technical Memorandum DRI-TM-14-01, March 2014 (acquired from info@dynres.com).

Smith, T.A., Koerner, J.E., Bieffe B12 Helmet Testing to AS/NZS 1698:2006, Dynamic Research, Inc. Technical Memorandum DRI-TM-14-67, 16 September 2014 (acquired from info@dynres.com).

Snook, C., An assessment of passive roll over protection for Quad Bikes, University of Southern Queensland, Australia, 2009.

Van Auken, R.M., and Zellner, J.W., Preliminary Analysis of the Effects of ATV ROPS on Rider Injury Potential, Second Revision, DRI-TR-96-4B, Dynamic Research, Inc., February 1997a.

Van Auken, R.M., and Zellner, J.W., Initial Analysis of the Effects of ATV Flexible ROPS on Rider Injury Potential, First Revision, DRI-TR-97-1A, Dynamic Research, Inc., March 1997b.

Van Auken, R.M., and Zellner, J.W., ATV ROPS/CPD Tests and Simulations, Volumes I - III, Dynamic Research, Inc. Technical Report DRI-TR-98-2, October 1998.

Van Ee, C., Toomey, D. and Morosky-Browne, B., "ATV Rollover, Operator Response, and Determinants of Injury: Implications for Crush Protection Devices," Design Research Engineering, Novi, Michigan, 14 November 2012.

Wordley, S. and Field, B., Quad Bike Safety Devices: A Snapshot Review, Department of Mechanical and Aerospace Engineering, Monash University, Victoria, Australia, 3 February 2012.

Wordley, S., Quad bike Crush Protection Devices (CPDs): updates to ISCRR Snapshot Review C-I-12-022, Department of Mechanical and Aerospace Engineering, Monash University, Clayton Campus, Victoria, Australia, 28 November 2012.

Zellner, J.W., Kebschull, S.A., Van Auken, R.M., Lenkeit, J.F. and Broen, P.C., Review and Analysis of MUARC Report "ATV Injuries and Deaths", and Additional Simulations and Initial Testing of MUARC ATV Rollover Protection System (ROPS), Dynamic Research Inc. Technical Report DRI-TR-04-01, Torrance, California, 2004.

Zellner, J.W., Van Auken, R.M., Kebschull, S.A., and Munoz, S., "Injury Risk-Benefit Analysis of Rollover Protection Systems (ROPS) for All-Terrain Vehicles (ATVs) Using Computer Simulation, Full-scale Testing and ISO 13232", Paper F2008-08-009, FISITA World Automotive Congress Proceedings, Volume II (Vehicle Safety), 14-19 September 2008, Munich, Germany, pp 14-27.

Zellner J.W. and Smith, T.A., Approximate AIS injury coding of 113 US/UK ATV Accidents, Dynamic Research Inc. Technical Report, DRI-TM-04-77, December 2004, and DRI-TM-04-77-1, revised November 2011.

Zellner, J.W., Kebschull, S.A., and Van Auken, R.M., Updated Injury Risk/Benefit Analysis of Quadbar Crush Protection Device (CPD) for All-Terrain Vehicles (ATVs), Dynamic Research, Inc. Technical Report TR-11-12, First Revision 6 August 2012, Second Revision 29 October 2014 (accessed at www.dri-atv-rops-research.com on 22 November 2014).

Zellner, J.W., Van Auken, R.M. and Kebschull, S.A., Replies to Lower (2011) Comments, Dynamic Research, Inc. Technical Report DRI-TR-12-05, Second Revision, 20 August 2012a (accessed at www.dri-atv-rops-research.com on 10 July 2013).

Zellner, J.W., Van Auken, R. M., and Kebschull, S.A., Comments on Monash/ISCRR Report by Wordley and Field (2012) Entitled "Quadbike Safety Devices: A Snapshot Review", Dynamic Research Inc. Technical Report DRI-TR-12-10, First Revision 17 October 2012b (accessed atwww.dri-atv-rops-research.com on 10 July 2013).

Zellner, J., Kebschull, S., and Van Auken, R., "Evaluation of Injury Risks and Benefits of a Crush Protection Device (CPD) for All-Terrain Vehicles (ATVs)," SAE Int. J. Passenger Cars – Electron. Electr. Syst. 7(1):2014, pp 41-72, doi: 10.4271/2013-32-9173.

Zellner, J.W., Kebschull, S.A., and Van Auken, R.M., "Predicted Effectiveness of Potential Injury Countermeasures for All-Terrain Vehicle Overturns", Paper No. 14-2272, The 93rd Annual Meeting Compendium of Papers, Transportation Research Board, Washington, D.C., January 12–16, 2014b.