ABANDON THE MOUSE RESEARCH SHIP? NOT JUST YET!

Marcin F. Osuchowski¹, Daniel G. Remick², James A. Lederer³, Charles H. Lang⁴, Ansgar O. Aasen⁵, Mayuki Aibiki⁶, Luciano C. Azevedo⁷, Soheyel Bahrami¹, Mihaly Boros⁸, Robert Cooney⁹, Salvatore Cuzzocrea¹⁰, Yong Jiang¹¹, Wolfgang G. Junger¹², Hiroyuki Hirasawa¹³, Richard S. Hotchkiss¹⁴, Xiang-An Li¹⁵, Peter Radermacher¹⁶, Heinz Redl¹, Reinaldo Salomao¹⁷, Amin Soebandrio¹⁸, Christoph Thiemermann¹⁹, Jean-Louis Vincent²⁰, Peter Ward²¹, Yong-Ming Yao²², Huang-Ping Yu²³, Basilia Zingarelli²⁴, and Irshad H. Chaudry²⁵,a

¹Ludwig Boltzmann Institute for Experimental and Clinical Traumatology in the AUVA Research Center, Vienna, Austria
²Boston University School of Medicine, Boston, MA, USA
³Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA
⁴Pennsylvania State College of Medicine, Hershey, PA, USA
⁵Oslo University Hospital, Oslo, Norway
⁶Ehime University, Matsuyama, Japan
⁷Research and Education Institute, Hospital Sirio-Libanes, São Paulo, Brazil
⁸Institute of Surgical Research, University of Szeged, Szeged, Hungary
⁹State University of New York Upstate Medical University, Syracuse, NY, USA
¹⁰Department of Biological and Environmental Sciences, University of Messina, Messina, Italy
¹¹Southern Medical University Guangzhou, Guangzhou, China
¹²Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA
¹³Chiba University Graduate School of Medicine, Chiba, Japan
¹⁴Washington University School of Medicine, St. Louis, MO, USA
¹⁵University of Kentucky College of Medicine, Lexington, KY, USA
¹⁶Ulm Medical University Clinic, Ulm, Germany
¹⁷Division of Infectious Diseases, Department of Medicine, Hospital São Paulo - Escola Paulista de Medicina, Federal University of São Paulo, Sao Paulo, Brazil

Corresponding Authors: Dr. Marcin F. Osuchowski Ludwig Boltzmann Institute for Experimental and Clinical Traumatology Donaaschengensstrasse 13 A-1200 Vienna, Austria Tel: +43-1-33110-469; Fax: +43-1-33110-460 marcin.osuchowski@trauma.lbg.ac.at Dr. Irshad H. Chaudry Center for Surgical Research University of Alabama at Birmingham G094 Volker Hall 1670 University Blvd. Birmingham, AL 35294-0019 Phone: 205-975-2195; Fax: 205-975-9719 ichaudry@uab.edu.

Competing interests: The authors declare that no competing interests exist.

“Extraordinary claims require extraordinary evidence.” Carl Sagan
Abstract

Many preclinical studies in critical care medicine and related disciplines rely on hypothesis-driven research in mice. The underlying premise posits that mice sufficiently emulate numerous pathophysiological alterations produced by trauma/sepsis and can serve as an experimental platform for answering clinically relevant questions. Recently the lay press severely criticized the translational relevance of mouse models in critical care medicine. A series of provocative editorials were elicited by a highly-publicized research report in the *Proceedings of the National Academy of Sciences* (*PNAS*; February 2013), which identified an unrecognized gene expression profile mismatch between human and murine leukocytes following burn/trauma/endotoxemia. Based on their data, the authors concluded that mouse models of trauma/inflammation are unsuitable for studying corresponding human conditions. We believe this conclusion was not justified. In conjunction with resulting negative commentary in the popular press, it can seriously jeopardize future basic research in critical care medicine. We will address some limitations of that *PNAS* report to provide a framework for discussing its conclusions and attempt to present a balanced summary of strengths/weaknesses of use of mouse models. While many investigators agree that animal research is a central component for improved patient outcomes, it is important to acknowledge known limitations in clinical translation from mouse to man. The scientific community is responsible to discuss valid limitations without over-interpretation. Hopefully a balanced view of the strengths/weaknesses of using animals for trauma/endotoxemia/critical care research will not result in hasty discount of the clear need for using animals to advance treatment of critically ill patients.

Keywords

mouse models of critical illness; trauma; endotoxemia; sepsis; burn

Introduction

In the USA (1) and European Union countries (2), approximately 15 and 7 million laboratory rodents, respectively, are used annually for research and testing. While it is difficult to precisely define the exact number of mice used in the field of critical care illnesses (such as trauma, burn, infections, endotoxemia, and sepsis), it is not an
overstatement that the majority of preclinical studies rely on this species. As of the end of October 2013, a search of PubMed produced between approximately 4200 and 24300 publications in response to a “mouse and burn/sepsis/shock/trauma” query (queries ordered from the lowest to highest number of hits). One of these studies, the collaborative report by Seok et al. published in the February 26, 2013 issue of the Proceedings of the National Academy of Sciences (PNAS), identified a potentially serious mismatch in the translational utility of the mouse-to-human data in the area of critical care medicine (3). The original data stem from the Inflammation and Host Response to Injury, Large Scale Collaborative Research Program Project (under the GLUE GRANT scientific consortium; www.gluegrant.org/index.htm). The aim of the GLUE GRANT project was to compare the genetic responses of humans and mice following burn, trauma and endotoxemia by analyzing approximately 5000 human genes and their mouse orthologs. Based on their interpretation of these results, the authors concluded that “genomic responses in mouse models poorly mimic human inflammatory diseases” and claimed the mouse gene profile response appeared random when compared to the human gene response to burn, trauma, or lipopolysaccharide (LPS) challenge. To those investigators familiar with how the immune system reacts to these innate stimuli, this broad generalization of their analysis appears, at least partly, unwarranted. Nevertheless, this labor-intensive study (3) is among a growing list of publications (4-6) challenging the merits of using mice and other animal models in basic and preclinical research and thus should be thoroughly discussed.

The Ripple Effect

Given the surprising and controversial nature of the data concerning mouse-to-human (in)compatibilities in tested inflammatory disease models, the findings of the PNAS paper were quickly publicized in the lay press. The initial account of the research in the New York Times entitled “Mice Fall Short as Test Subjects for Some of Humans’ Deadly Ills” (7) led to a subsequent ripple effect in the form of several alarming follow-up editorials, posts and/or blogs (8-11). Their collective conclusion was clear and implied that decades of mouse-based research culminated in few scientific advances, wasted precious research opportunities, and was a poor use of taxpayers’ money. Consequently, given that “mouse models of inflammation are basically worthless” (10), “it seems that researchers have tortured mice in vain for decades in the search for drugs to help humans recover from certain traumas, like severe burns, blunt force, and sepsis” (11). There is concern that the sensational tone of those communications will be damaging to preclinical mouse-based research programs. As a result, public perception, research progress and funding support for basic discovery and hypothesis-driven research for many medical disciplines may be impeded. As the authors of the original PNAS paper were mostly directing their criticisms towards inflammation, trauma, shock, and sepsis research, we felt compelled, following the recent comments by others (12-16), to collectively address their controversial conclusions. By discussing its main limitations, we aim to delineate the boundaries within which the work of Seok et al. should be viewed and evaluated. Importantly, we also provide objective information demonstrating that animal research using mice has led to ground-breaking studies that have improved patient care and outcomes.

Shock. Author manuscript; available in PMC 2015 June 01.
Lost in Translation: What Does the PNAS Study Really Say?

Seok and colleagues report that the genomic response to trauma, burns, or endotoxin challenge shows an extremely low correlation between mice and humans, while these different types of injury responses showed high similarity among humans. The authors state in the first paragraph that “Among genes changed significantly in humans, the murine orthologs are close to random in matching their human counterparts (e.g., R² between 0.0 and 0.1).” We contend that the authors have over-interpreted their data due to the many limitations of their study design and analysis, some of which they have failed to acknowledge. Furthermore, it remains uncertain whether and/or to what extent the results of gene expression profiling should be used to judge the biological validity of animal models for human disease. Although we disagree with the overall conclusion and interpretation of this PNAS report, the intention of this manuscript is not to lessen the value of the authors’ work but to analyze conclusions of the study in an appropriate evidence-based framework. The following is a partial list of limitations and issues that were identified subsequent to the publication of the PNAS paper (17-19) and that were featured in a debate session at the 2013 annual Shock Society meeting in San Diego, CA, USA (20).

1) Comparing strain, sex and age

Gene profiles of a highly heterogeneous (outbred) population of burn/trauma/endotoxemic male and female patients were compared to inbred, genetically identical C57BL/6J male mice at the approximate age of 2 months. Using inbred mice that are genetically identical for such a comparative analysis is equivalent to comparing a single individual burn or trauma patient to 167 trauma or 244 burn patients. Furthermore, comparing individual responses to these injuries in inbred versus outbred subjects represents an important study limitation because of the differential immune system response among inbred mice to various parasitic (21;22), viral (23;24) and bacterial (25;26) infections. The most recent review by Fink (27) offers a deeper insight into the limitations of the inbred strains in modeling of sepsis (also in the context of the PNAS article). Summarizing, testing a single strain of mice does not justify the assertion that all mouse models poorly mimic human inflammatory diseases given that this or any other strain fails to represent the genetic diversity of the entire mouse population. Furthermore, comparison of a single mouse sex to the mixed male/female population of human patients demands further caution regarding data interpretation given that sex modifies leukocyte responses after trauma-hemorrhage (28-30), bacterial infection (31) or LPS (32;33). Additionally, it cannot be assumed that 8-week-old mice (used in the PNAS study) constitute a reliable surrogate for older patients. In the C57BL/6J strain, the age of 8 weeks corresponds approximately to the prescent human age of 8 years (34). This contrasts sharply with the average age of 34 years in trauma patients (35), 30 years in healthy control subjects and 18-40 years in the eight volunteers injected with endotoxin (3;35). It remains to be established how strongly (if at all) this particular age disparity can influence the studied responses in mice and humans. However, the role of age should not be minimized as differential age-dependent outcomes and responses in the immuno-inflammatory compartment were demonstrated in pediatric versus adult patients with trauma (36; 37), burns (38;39), endotoxemia (40) and infections (41;42) as well as in corresponding mouse models (43-46). Of note, although not directly pertinent to data analyzed in the study

Shock. Author manuscript; available in PMC 2015 June 01.
by Seok and colleagues (3) (as leukocytes from septic patients were not analyzed), age is a major mismatch in the mouse-to-human comparison of data from sepsis syndromes; while a majority of human septic patients are old, preclinical sepsis models (including the most relevant) typically rely on young mice (47).

2) Gene expression analysis using unfractionated blood leukocytes

The composition of circulating leukocytes differs markedly between humans and mice (i.e., 60% versus 15% neutrophils and 30% versus 70% lymphocytes, respectively) (6) and a recent study by Shay et al. demonstrated that human versus mouse granulocytes and lymphocytes have distinct gene expression profiles (48). Because the comparison was made on only unfractionated whole blood leukocytes in the study, this translates into the profiling of a neutrophil-rich versus lymphocyte-rich sample, which undoubtedly skewed the final results. It must be noted that an oral communication by one of the PNAS co-authors indicated that reanalysis of genomic responses in a more narrow PMN population did not markedly improve the overall correlation (20), although a peer-reviewed publication of such a reanalysis has not yet been reported. The risks of imprecise comparison of genomic responses in circulating white blood cells (WBC) are not solely restricted to mouse versus human studies; earlier GLUE GRANT-based reports that analyzed various inflammation/injury scenarios using only mouse models, voiced identical concerns (49;50). Analyses of gene expression in discrete leukocyte subsets are certainly warranted as they provide more precise information regarding individual activation patterns in injury and/or infection.

3) Model and severity mismatch

The influence of this potential limitation was suggested by the authors in the original paper. For example, mice may have higher resistance to inflammation/trauma/infection because they are housed in a controlled environment and most often lack predisposing conditions present in the human population. This injury severity mismatch may translate into a dissimilar response both on the genomic and protein levels. Though animal models are designed to produce a pronounced, clinically relevant post-traumatic physiological effect, a near-death severity threshold is often needed since milder severity only recapitulates some features of trauma or sepsis response (51-53). It is also noteworthy that due to their innate high resistance to trauma and inflammation, mouse models cannot adequately recapitulate a full pallet of the most severe responses that occur in patients. This may be partly due to yet another mismatch: it is generally appreciated that any patient admitted to the ICU and survives would have likely died without intervention. Our inability to replicate this aggravating phenomenon may be another influential limitation in animal modeling.

To approximate sepsis, the authors used LPS delivery into human volunteers and mice for their gene expression profile comparison. Although this is a commonly used model in humans that supposedly mimics some clinical features of the sepsis syndrome, it is not a true model for sepsis. Sepsis and sepsis syndrome are complex responses and most of the existing experimental models fail to reproduce the entire spectrum of sepsis syndromes diagnosed in patients, especially the bolus injection of LPS. In other words, it is inappropriate to compare aged patients suffering from monobacterial pneumonia induced bacteremia with the cecal ligation and puncture (CLP) model of polymicrobial peritonitis.

*Shock. Author manuscript; available in PMC 2015 June 01.*
Other mismatched secondary comparisons were made based on the Gene Expression Omnibus database (54); data from each of the existing specific sepsis models should only be compared to the corresponding clinical condition (47). Another debatable point, recently raised in a letter by Cauwels et al. (12), is that the relative dose of LPS differed markedly between mice and human volunteers. It is also important to consider that besides genetic background, environment and underlying diseases can markedly affect the host response to infection. In such a scenario mice can be rendered either hyporesponsive or hyperresponsive to LPS with major differences in inflammatory responses that in turn alter susceptibility to infection. For example, mice with chronic kidney disease exhibit an increase in morbidity when subjected to sepsis (55). Increased inflammatory responses may be induced by previous Propionibacterium acnes infection, hepatotoxic agents (e.g., D-Galactosamine) and growing Lewis lung carcinoma, while exposure to minute amounts of LPS renders the animals tolerant to LPS (56). This modulation is likely to occur in humans as well, in whom underlying diseases such as diabetes mellitus and end stage renal dysfunction are known to impair host response to infection (47). Thus, diverse underlying conditions or exposure to pathogens may further underscore differences between experimental studies with healthy animals and studies in humans in clinical settings.

4) Analysis of responses in a single source

Many critically ill patients trigger activation of virtually all body systems so that the total amount of secreted cytokines comes from multiple cells and organs (Fig. 1). Due to technical and/or diagnostic ease, blood, with its cellular and soluble components, is the most frequently used source of information in critical conditions (57-60). Although the blood-based diagnostics have proven utility in many facets of ICU monitoring, blood leukocytes represent a relatively narrow source considering the entire pool of inflammatory mediators released to the systemic circulation after trauma and during severe infections (61;62). Given that virtually all inflammatory mediators detected in the blood are produced by various sources (Fig. 1), it cannot be ignored that the interspecies compatibility is better when other sites of cytokine synthesis are compared (e.g., hepatocytes). Yet, despite rapid technological advancements in the ICU, such interspecies comparisons are currently either beyond reach or very difficult to perform. While this caveat is not a shortcoming of the paper, it does highlight the fact that the systemic or organ-specific contribution to the inflammatory response needs to be taken into consideration when interpreting results.

5) Compartmentalization of the immune-inflammatory response

The multi-source synthesis of inflammatory mediators has yet another important aspect – the compartmentalization of the immuno-inflammatory response. It has been suggested that in systemic inflammatory response syndrome (SIRS) and/or infection, compartmentalized synthesis and release of inflammatory cytokines are equally important (63-65) and that cells other than leukocytes may be responsible for morbidity and mortality in trauma (61), inflammatory shock (66), endotoxemia (67-69) and/or infection (70). Studies have also shown that peripheral blood cells fail to reflect what occurs in the tissue fixed cells within the same or different organs (30;49;50;71;72). Furthermore, the concept of compartmentalization pertains to coexistence of differential (and often contrasting) responses that depend on the specific location of the immune-competent machinery and this
notion has been supported by numerous preclinical studies. For example, in mice, hemorrhage resulted in a contrasting cytotoxic capacity with reduction in peritoneal and splenic macrophages and enhancement in Kupffer cells (73). The disparity in the post-hemorrhage response also extends to the cytokine compartment: transcriptional activity of TNF-α, IL-1β and TGF-β (but not IL-6) as well as the release capacity of TNF-α and IL-1 were shown to be enhanced in Kupffer cells but diminished in peritoneal and splenic macrophages (73;74). Remarkably, compartmentalization of the gene and protein expression response after trauma-hemorrhage (30;75), infection (71) and no challenge (76) was also reported to occur in different cell types of the same organ (e.g., liver) and in cells of the same origin (e.g. mononuclear cells) but different locations (i.e., blood, spleen, peritoneum). Interestingly, data reported by the two aforementioned GLUE GRANT studies that compared genomic WBC and splenic responses in the mouse injury/infection model, further substantiate the notion of compartmentalization given that responses in those two compartments were highly dissimilar both at the early (50) and late phase after challenge (49). Collectively, the existing evidence strongly suggests that changes occurring in the blood compartment are not tantamount to the alterations arising in tissues/organs. Thus, any changes in treatment protocols based on changes occurring in circulating leukocytes only, regardless of whether on the genomic or protein level, and without considering the role of less diagnostically accessible compartments, could be very detrimental to patients.

6) Genome-to-protein gap

The poor correlation between mouse and man of the genomic responses reported in the original PNAS publication was not verified on the protein level. While fully understandable because of the labor-intensive nature of genomic screening, the wide gap between the genomic activation and the final protein product should not be summarily discounted. We speculate, for example, that correlation of circulating cytokines between patients and corresponding mouse models, had it been performed, would have shown better correlation than the changes in mRNA content in leukocytes. Prior work by the GLUE GRANT investigators (77) showed that using the same type of endotoxin in mice and humans produced nearly identical kinetics of cytokine production. Additionally, the concentrations of many of the cytokines or cytokine inhibitors that were produced were relatively close. Also, both mice and humans showed equivalent lymphopenia after endotoxin injection. Furthermore, the existing evidence demonstrates comparable temporal patterns for systemic inflammatory responses after both injury (78;79) and/or systemic infection (47), especially when patients are matched with the corresponding (and appropriate) mouse models. Collectively, these data suggest the total concentration of circulating cytokines is either substantially enriched by non-leukocyte sources (Fig. 1) or strongly influenced by various post-transcriptional/translational modifications (Fig. 2). Regarding the latter point, the strong relationship between mRNA and protein abundance levels is not self-evident and cancer research demonstrates that the correlation greatly varies. For example, comparison of mRNA and protein expression of three genes/protein pairs (MMP-2, MMP-9 and TIMP-1) showed no significant relationship in prostate cancer patients (80), while in human lung adenocarcinomas, only 20% of the 98 total screened genes demonstrated statistically significant correlation with their respective products (81). In human bladder cancers (82), gene-to-protein correlation was highly significant in some targets and low to negligible in...
others. A very similar outcome (i.e., high correlation variability between protein and mRNA expression levels) was demonstrated in circulating monocytes from healthy females (83) and human livers (84) suggesting that the use of mRNA expression to predict protein expression levels even in healthy organisms appears to be burdened with a relatively high uncertainty. The presence of a severe condition such as trauma or sepsis may further influence the gene-to-protein conduit. Overall, our current understanding of the dynamic chain of events occurring between genomic activation and arrival of the final product in diseased organisms remains poor and a straightforward gene-to-protein relationship should not be reflexively assumed in any biological system (85).

7) (Mis)matching the temporal response patterns

The analysis of the genomic data included comparisons of temporal expression patterns between human and mouse models. This is a commendable approach given the rapid fluctuation produced by inflammatory responses in trauma (blood loss and/or burn injury) and endotoxemia (47;86;87). Yet, execution of such a longitudinal comparison with large databases is particularly prone to both type I and II errors. One example of such a potential inaccuracy was discussed in the most recent commentary by Perlman et al. (14): the authors suggested that temporal dynamics of mouse and human critical conditions should not be matched 1:1 on the time scale given that compared to patients, murine disease models evolve over a shorter time course. A similar concern regarding the human-to-mouse time mismatch was echoed by Osterburg et al. (15). Furthermore, preliminary re-assessment of the same GLUE GRANT data, with a specific focus on the temporal match, by another independent group (ImmGen Consortium) was presented at the recent Shock Society Pro and Con Debate (20). This alternative analysis demonstrated markedly better human-to-mouse correlations (e.g., reaching r = 0.41 on day 7 post-burn injury by Fold-Change Quadrant Plot analysis) suggesting the original analysis of temporal responses might not have been sufficiently rigorous. The complete and detailed findings of the aforementioned reassessment will be soon submitted for peer review and will be published in a separate report to allow impartial comparison of the two different analytical methods applied.

In summary, the above concerns warrant additional mouse versus man studies and verification by an independent investigative team to validate the analysis of the findings reported by Seok and coworkers. If confirmed, the asserted dissimilarity of the genomic responses in human versus mouse leukocytes will have important and long-term implications for preclinical animal research, for example, alerting against the inter-species translational incompatibility of data generated in preclinical (i.e., in the mouse) testing of potential therapeutic genetic manipulations on the level of circulating leukocytes. At present, however, due to the rapid character of interventions in critical medicine, the vast majority of inflammation-targeted therapies focus on circulating inflammatory mediators – irrespective of the site of their synthesis, and focus less on manipulations of gene expression.

The Good and the Bad: The Role of Mouse Models in Critical Illness

Compared to other protracted disease conditions such as cancer or atherosclerosis, critical care research faces more difficult hurdles. These challenges stem from the fact that critical illness typically progresses very quickly, triggers rapid reactions from all body systems,
results in heterogeneous reactions and causes extreme fluctuations in the elicited responses. Consequently, the challenge of trying to develop clinically relevant mouse models to simultaneously mimic all components of human critical injury and illness likely represents a fool’s errand. Paradoxically, the wide public uproar provoked by the work of Seok and colleagues may simultaneously be - regardless of the ultimate accuracy the study’s findings - its great service to the scientific community as it has elevated the discussion on the adequacy of preclinical mouse modeling in the contemporary research realm to a deserving premiere spot. In doing so, it is likely to also stimulate the discussion of the fidelity of mouse models designed to mimic other human pathological conditions (e.g., heart disease, obesity, cancer, etc.).

The limitations of existing mouse models of critical care illnesses (and beyond) are not a new concern. Apart from the ones discussed here, previous papers have focused on the mouse-to-human mismatch, e.g., characterizing major differences in the immuno-inflammatory signature (6). Additionally, responses to inflammatory stimuli (88;89), biochemical functionality of homologous proteins (90), makeup of serum proteins (91) as well as the influence of so called “cold stress” in the laboratory setting (92) have been identified as important differences between mice and humans. However, some of these same arguments could be used to criticize the way in which clinical studies are conducted. For example, laboratory isolation approaches for blood leukocytes vary widely between laboratories. This could lead to contrasting results in experiments using patient samples in a manner similar to that of mouse studies. Also, most human studies are limited to using blood samples, while other tissue and organ compartments can be effectively studied in mice, which improves the scientific profundity of well-designed animal research studies. No less important for modeling/translational purposes are variations within populations of human patients themselves. Trauma/sepsis patients typically receive allogeneic blood transfusion (versus shed blood in animals), different doses of morphine-based products (versus lack of and/or set doses of non-/morphine substances), inotropic agents and others, a majority of which produce strong immune-inflammatory effects that are very difficult to account for and match experimentally (93). Thus, patient studies are fraught with numerous small and large differences in the manner in which clinical care and standard procedures are carried out at different institutions. Furthermore, apart from the biologically-based differences, preclinical testing is simultaneously influenced by no less important study design flaws such as: 1) an excessive focus on early (acute) events; 2) age mismatch; 3) a lack of comorbidities; 4) using pre-treatment versus post-treatment in an appropriate manner; 5) lack of broad spectrum antibiotic coverage for sepsis treatment studies; and 6) difficulty in reproducing specific ICU conditions in animal models (47). This is far from an exhaustive list of mouse-to-human mismatches and/or potential confounding factors - many new pieces of this puzzle are yet to be identified.

From a global perspective, however, these differences do not appear to supersede similarities as there are countless physiological and pathological traits shared by both species in response to critical illness. We have selected dozens of relevant examples demonstrating the striking similarities in the responses of mice and humans (Table 1). This list could serve as a starting point for designing animal models for answering clinically
relevant questions. Ideally, the origins of a “perfect” animal model should be deeply rooted in the clinical problem-solving and the model itself should reproduce as closely as possible the entire spectrum of pathophysiological consequences and the mechanisms of the human condition it aims to duplicate. Such an ideal match may be challenging to achieve, not only with mice or rats but with larger species as well, even non-human primates. An obvious first step in responsible modeling is to identify and verify a murine system with an acceptable resemblance to the studied critical illness and in the specific context of the defined scientific question. For example, one should not set out to study septic acute lung injury in CD-1 mice as this strain does not typically develop this condition (94), or to expect that pre-treatment data generated in a burn model conducted in healthy 4-week-old male mice will be translatable to the entire spectrum of human patients. The initial selection of the model must be subsequently followed by a fine tuning of the study design and, finally, a realistic interpretation of the acquired data. The latter element requires careful consideration. Too frequently (the PNAS article notwithstanding) lax interpretations of animal-based results are published in the scientific literature creating confusing (if not outright misleading) “background noise” (95). Compulsory disclosure of relevant animal experimentation details in compliance with the ARRIVE Guidelines (http://www.nc3rs.org.uk; recently adapted by the Shock journal; Instructions for Authors. Shock 41(1), January 2014) will partly help to remedy this confusion. A more diligent peer-review process is even stronger medicine given the growing stock market-like competition in the area of scientific publishing (96).

Achieving the goal of implementing the elements listed above will likely generate datasets that are much more precise and relevant to patients. Yet, old habits die hard and many misconceptions surrounding applicability of data from mouse models are likely to have long half-lives. The most recent exchange between Cauwels *et al.* (12) and the authors of the original paper (3) is perhaps a telling example: in the reply letter to Cauwels *et al.* (97), the fact that anti-TNF treatment was life-saving in mice administered a lethal dose of LPS (98) but failed in septic patients, supported their notion that mouse models are generally unfit to predict human inflammatory diseases such as sepsis. However, when anti-TNF antibodies were used in a clinically relevant model of sepsis (i.e. cecal ligation and puncture) (99;100), they also failed to have any protective effects against sepsis (101); a finding similar to the lack of efficacy seen in patients. Thus, it is humans, and not mice, with their incomplete understanding of sepsis pathophysiology coupled with the use of an inappropriate animal model, which are to be blamed for the spectacular collapse of anti-inflammatory sepsis trials. After “successfully” executing various treatments (i.e., against circulating cytokines or endotoxins) in the mouse (98), rat (102), rabbit (103), dog (104) and non-human primates (105), the striking failure of similar anti-inflammatory therapeutic protocols in septic patients has brought a painful realization that injecting mice and other species with a lethal dose of bacterial LPS is not a good predictive model for a typical human sepsis. It must be stressed, however, that a complete renunciation of anti-inflammatory treatments based on the failed trials would be equally short sighted and the most recent meta-analysis argues that application of anti-TNF agents in septic patients should be revisited (106). It has become clear that responses elicited in sepsis are highly mixed and the immunosuppressive component frequently exceeds hyperinflammation (107). Hence, in smaller cohorts of patients treated based on the similar type of acquired sepsis (e.g. meningococcal septicemia),

*Shock*. Author manuscript; available in PMC 2015 June 01.
their well-defined immuno-inflammatory makeup and/or predicted susceptibility toward the tailored treatment, even the “notorious” anti-TNF intervention may be lifesaving (108). Limited evidence in support of implementing such tactics is already available (109;110; www.clinicaltrials.gov identifier: NCT01046669).

Clearly, sepsis and septic shock are complex disease processes. We suggest multiple animal models should be considered for basic research and preclinical testing of therapeutics. The remarkable capability of contemporary science has opened many new investigative avenues such as emergence of humanized mice (111, 112, 113) and access to a growing selection of recombinant inbred mouse strains (with specific defined genetic traits) from the Collaborative Cross Project (114; 115). Yet, whereas using inbred mice for basic research remains important to advancing knowledge and may allow much more nuanced investigation of gene environment and gene-pathogen interactions, the efficacy testing of beneficial treatments should be tested in outbred mice as well. Attempts to test the efficacy of treatments with multiple types of infectious disease models (e.g., CLP, pneumonia, urosepsis etc.) should be considered and tested before moving forward with efficacy trials in large animals and then patients.

**Final Conclusions**

The current reality is that other than whole blood assays and isolated single organ cultures, animal models are the only viable and fully intact biological systems which allow examination of clinically relevant hypotheses and studies to decipher underlying mechanisms of biological phenomena. Mice have been in the forefront of these investigations precisely for the reason that they have served as an origin of many subsequent successes at the patient's bedside, including the area of critical care medicine. Yet, it is evident that any mouse study is merely a beginning in the long process that requires caution and series of subsequent verification steps. To continue with in vivo research in an ethically responsible way, the scientific community has an obligation to seek improvements and implement more fitting solutions in the currently used mouse model systems so they continuously adapt to the evolving understanding of the respective human critical illnesses and not vice versa. We must remain cognizant of the known limitations of the models, share newly discovered incompatibilities and be willing to abandon the erroneous models if necessary. In relation to this discussion on the relevance of mouse models in critical care medicine, the allegory of inflammatory response appears to be very fitting. Specifically, an exaggerated or excessively weak inflammatory reaction will typically lead to a poor outcome in an ICU patient. In a similar manner, excessive trust or hasty discounting the usefulness of mouse models for research will have a negative impact on preclinical critical care research and ultimately result in fewer discoveries which improve patient care and outcome.

**Acknowledgments**

We gratefully acknowledge Jean-Marc Cavaillon for his constructive input regarding the merits and balance of the manuscript. We thank Pia Rademann for her kind help with the Reference Manager and in drawing the Figures.
Reference List


9. Love, SM. We are not mice. 2013-02-14 (Huffingtonpost.com) Available at: http://www.huffingtonpost.com/susan-m-love/research-mice_b_2680201.html

10. Derek, Low. Mouse Models of Inflammation Are Basically Worthless. Now We Know. 2013-02-12 (pipeline.corante.com) Available at: http://pipeline.corante.com/archives/2013/02/13/mouse_models_of_inflammation_are_basically_worthless_now_we_know.php


Shock. Author manuscript; available in PMC 2015 June 01.


Shock. Author manuscript; available in PMC 2015 June 01.


Fig. 1. Sources of inflammatory mediators produced in response to a critical condition
A typical critical illness, regardless whether of purely traumatic or infectious origin, triggers a broad inflammatory response that involves virtually all organ and cellular systems. The sum of inflammatory mediators circulating in the blood, the most expedient source of diagnostic information, is composed of proteins synthetized at multiple sites and sources. The simplified (given the overlap of cells and organs) schematic displays the most important and recognized sources of cytokine production. Black framing indicates the organ/cell populations which were used in the comparative genomic response study by Seok et al. (3). CNS: central nervous system.
Fig. 2. From DNA to the final product - Regulation of gene expression
The path from the genomic activation to the final protein is interrupted by a series of complex regulatory steps that may either completely halt and/or markedly modify the abundance of the final product. The simplified schematic lists the most important regulatory mechanisms between DNA activation and synthesis of the ready protein. 1) To initiate transcription, demethylated DNA must be bound to acetylated histones; 2) RNA-polymerase and transcription factors bind on DNA to start transcription; 3) mRNA leaves the nucleus and is translated into protein; 4) Some proteins are produced in a dormant state and require subsequent activation. Green arrow indicates the first regulatory step between genomic activation and generation of the corresponding mRNA. Red arrows indicate remaining regulatory steps prior to the emergence of the final protein coded by the mRNA.
Table 1

<table>
<thead>
<tr>
<th>#</th>
<th>Translational phenomenon/response</th>
<th>Specific comments - mouse</th>
<th>Specific comments - human</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Antibodies to TNF&lt;sup&gt;α&lt;/sup&gt; given indiscriminately fail to reduce sepsis mortality</td>
<td>BALB/c mice were pretreated with antibodies to TNF prior to CLP sepsis. The murine studies were published 3 years before the failed human trials (101;116)</td>
<td>Anti-TNF antibodies failed to be an effective treatment strategy in a general population of septic patients (117;118)</td>
</tr>
<tr>
<td>2</td>
<td>Pretreatment with an anti-TNF strategy prevents early systemic inflammatory response syndrome (SIRS)</td>
<td>Passive immunization with the antiserum to TNF-α in BALB/c mice protected them against the lethal hyperinflammation by Escherichia coli LPS (98)</td>
<td>Anti-TNF&lt;sub&gt;α&lt;/sub&gt; therapy was effective in humans with louse-borne relapsing fever when given as a pre-treatment against Jarisch-Herxheimer reactions (119)</td>
</tr>
<tr>
<td>3</td>
<td>Low-dose steroid therapy is associated with decreased mortality in septic mice and humans</td>
<td>Demonstrated in C57BL/6 male mice subjected to CLP and treated with different corticoid concentrations; low but not high-dose steroids improved 21 day survival (120)</td>
<td>Early initiation of low-dose corticosteroid therapy decreased mortality in septic shock patients (121)</td>
</tr>
<tr>
<td>4</td>
<td>Regulation of chemotactic behavior of mouse and human neutrophils via purinergic signaling</td>
<td>Human and mouse neutrophils rely on same purinergic receptor subtypes (P2Y2, A3, and A2a receptors) for autocrine signaling (122;123;124)</td>
<td>Demonstrated &lt;em&gt;in vitro and in vivo&lt;/em&gt;; mice are suitable to study chemotaxis in inflammation, trauma, and sepsis (122;123;124; NCT01180361&lt;sup&gt;b&lt;/sup&gt;)</td>
</tr>
<tr>
<td>5</td>
<td>Human and mouse neutrophils rely on similar signaling mechanisms for their activation during bacteria-induced acute lung injury</td>
<td>Increased nuclear activation of NF-kB in pulmonary neutrophils of mice after &lt;em&gt;in vivo&lt;/em&gt; administration with endotoxin (125;126)</td>
<td>Increased nuclear accumulation of NF-kB in peripheral or pulmonary neutrophils of human volunteers after &lt;em&gt;in vitro&lt;/em&gt; or &lt;em&gt;in vivo&lt;/em&gt; stimulation with endotoxin (127) or in peripheral neutrophils of patients with sepsis (128)</td>
</tr>
<tr>
<td>6</td>
<td>Sepsis always in MARS: simultaneous systemic release of both pro- and anti-inflammatory cytokines in sepsis</td>
<td>Demonstrated in ICR/CD-1 (outbred) female mice subjected to CLP sepsis (129;130)</td>
<td>Demonstrated in septic shock patients (131) and patients with post-operative abdominal sepsis (132)</td>
</tr>
<tr>
<td>7</td>
<td>IL-6 serves as a biomarker for sepsis mortality</td>
<td>IL-6 measured 6 hours after the onset of CLP sepsis in BALB/c (133) and CD-1 mice (129) accurately predicts survival</td>
<td>Patients with high levels of IL-6 are at increased risk of dying from sepsis (134;135)</td>
</tr>
<tr>
<td>8</td>
<td>Role of nicotinic receptors in inflammatory responses after endotoxemia is similar in mice and humans</td>
<td>Demonstrated in C57BL/6 mice and &lt;i&gt;c&lt;/i&gt;7 nicotinic receptor deficient mice; endotoxin-induced response was abrogated via activation of anti-inflammatory cholinergic pathway (vagus nerve stimulation) (136)</td>
<td>Human volunteers were administered endotoxin and GTS-21 (&lt;i&gt;c&lt;/i&gt;7nAChR agonist) or placebo to study anti-inflammatory effects of cholinergic pathway (137; NCT00783068&lt;sup&gt;b&lt;/sup&gt;)</td>
</tr>
<tr>
<td>9</td>
<td>Similar mode of pathogen-associated molecular patterns detection via toll like receptors (TLR) in mice and humans</td>
<td>TLR-4 was identified as the receptor that senses LPS in experiments with congenic sensitive (C3HHeN; C57BL/10ScSn) and resistant (C3H/HeJ and C57BL/10ScCr) mice (138); TLR-4 expression level determines the degree of LPS-susceptibility in mice (139)</td>
<td>Human volunteers administered with LPS demonstrated altered TLR-induced genes expression (140). TLR-signaling pathways are strongly modulated in septic patients (141)</td>
</tr>
<tr>
<td>10</td>
<td>Sepsis induces profound apoptosis of immune and gastrointestinal epithelial (GIE) cells</td>
<td>Demonstrated in CLP female ND4 mice (142) and &lt;i&gt;Pseudomonas aeruginosa&lt;/i&gt; pneumonia-induced septic FVB/N mice (143); apoptosis in B and T lymphocytes and dendritic cells. GIE cell apoptosis in large and small intestine</td>
<td>Demonstrated in patients who died of sepsis and sepsis and MODS; Data obtained by retrospective (rapid autopsy) and prospective (tissue resection) examination (144;145;146)</td>
</tr>
<tr>
<td>11</td>
<td>Sepsis increases while septic shock decreases the rate of hepatic gluconeogenesis</td>
<td>Demonstrated in mechanically ventilated C57BL/6 (inbred) male mice subjected to CLP sepsis (147)</td>
<td>Demonstrated in “infected” patients after surgery or trauma (148) and “bacteremic”/“complicated bacteremic” burn patients (149)</td>
</tr>
<tr>
<td>12</td>
<td>Increased nitric oxide reduces systolic contractility but supports (adaptive) left ventricular diastolic relaxation</td>
<td>Demonstrated in mechanically ventilated male C57BL/6 mice subjected to CLP sepsis (150)</td>
<td>Demonstrated in patients with septic shock (151), rat cardiomyocytes exposed to serum from patients with septic shock (152), and patients with chest pain (153)</td>
</tr>
<tr>
<td>#</td>
<td>Translational phenomenon/response</td>
<td>Specific comments - mouse:</td>
<td>Specific comments - human:</td>
</tr>
<tr>
<td>----</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>13</td>
<td>Identical pathomechanism of increased intestinal mucosal permeability during inflammatory conditions in mice and humans</td>
<td>Increased intestinal permeability is associated with an increase in IL-1β-induced NFκB activation and MLCK expression (154). This action involves p38 kinase and ATF-2 activation in mice and humans (155)</td>
<td>Production of IL-1β in ulcerative and Crohn’s colitis (156) increases intestinal epithelial tight junction permeability (157); these actions are mediated by an NFκB-dependent increase in MLCK gene transcription (158)</td>
</tr>
<tr>
<td>14</td>
<td>Epithelial tight junction barrier failure in mice and humans</td>
<td>Demonstrated in anti-CD3 murine diarrhea (159;160) model, Card15/NOD2- (159) and NOD2-deficient (160) mice; MLCK activation is necessary for epithelial barrier dysfunction and mucosal permeability increase (159, 160); NOD2 gene regulates this response (161,162)</td>
<td>Demonstrated in human colonic epithelial cell line HT-29/B6 (163), Caco-2 cells (164,165) and intestinal epithelial cells (165); TNFs impairs mucosal barrier function of the epithelial tight junction (163) by NFκB activation (164) and increasing gene/protein expression of MLCK (165)</td>
</tr>
<tr>
<td>15</td>
<td>Intrauterine group A streptococcal infections in mice and humans are similarly modulated by prostaglandin (PG) E2</td>
<td>Demonstrated in a mouse (C57BL6) GAS (Group A Streptococcus) infection model; PGE2 impaired the phagocytic ability of mouse peritoneal macrophages in vitro (166)</td>
<td>An in vitro human macrophage-GAS interaction model was used. In GAS infected human THP-1 (macrophage-like) cells, PGE2 impairs the phagocytic ability of THP-1 and of human placental macrophages (166)</td>
</tr>
<tr>
<td>16</td>
<td>LPS-tolerant cells present a reprogramming of gene expression and function following a second challenge with LPS</td>
<td>Differentially regulated genes were found in bone marrow-derived macrophages tolerant to LPS. The tolerizable (T; proinflammatory action) genes were typically not reinduced in macrophages upon second LPS exposure, while the nontolerizable (NT; antimicrobial action) genes were reinduced (167)</td>
<td>Demonstrated in human monocytes made tolerant to LPS (168,169); Class T genes encompass antigen presentation-related and pro-inflammatory cytokine genes; NT genes code for anti-inflammatory/microbial factors (168); T and NT genes modulate TLR signaling pathway (169)</td>
</tr>
<tr>
<td>17</td>
<td>Sepsis increases muscle protein degradation and proteasome activity</td>
<td>CLP-induced sepsis in B6 mice increased total and myofibrillar protein breakdown in EDL, with evidence of increased proteasome activity (170)</td>
<td>Critically ill patients with sepsis show increased muscle proteolysis and proteasome proteolytic activity (171)</td>
</tr>
<tr>
<td>18</td>
<td>Sepsis decreases lean body mass and increases muscle wasting</td>
<td>CLP-induced sepsis in C57BL/6 male mice decreases body weight, lean body mass and muscle mass (172)</td>
<td>Patients with severe peritonitis sepsis show loss of body weight and muscle mass and protein (173)</td>
</tr>
<tr>
<td>19</td>
<td>Sepsis induces elevated thrombin-antithrombin complexes are a part of TF/FVIIa pathway activation of coagulation</td>
<td>Using a mouse model peritonitis was induced by an intraperitoneal injection of live Escherichia coli (174)</td>
<td>Demonstrated in patients with severe bacterial peritonitis (175) and in healthy volunteers injected with endotoxin (176)</td>
</tr>
<tr>
<td>20</td>
<td>Sepsis induces systemic elevation of matrix metalloproteinase (MMP)-8 as part of the inflammatory process in humans and mice</td>
<td>Increased activity of MMP-8 was observed in plasma of C57BL/6 mice. Generation of MMP-8 genetically deficient mice was used to understand the biological function of the enzyme in inflammation (177)</td>
<td>Increased mRNA expression and activity of MMP-8 was observed in blood of pediatric septic patients and correlated with disease severity and mortality (177)</td>
</tr>
<tr>
<td>21</td>
<td>Similar mechanisms of mitochondrial dysfunction and mitochondrial biogenesis during sepsis in mice and humans</td>
<td>Mitochondrial dysfunction was demonstrated in CLP male BALB/c mice (178); mitochondrial biogenesis was associated with an increase in PGC-1α in peritoneitis by fibrin clot in male C57BL/6 mice (179)</td>
<td>Mitochondrial dysfunction was demonstrated in muscle biopsies of septic shock patients (180). Septic survivors presented increased expression of the mitochondrial biogenesis-related PGC-1α (181)</td>
</tr>
<tr>
<td>22</td>
<td>Inflammation-induced autophagy in tissues</td>
<td>Demonstrated in C57BL/6 male mice subjected to CLP (182) and Apg16L1 deficient healthy mice (183) and mice infected with Salmonella enterica (184); autophagy regulated by Apg16L1 gene</td>
<td>Demonstrated in patients that died of sepsis (182) and in patients with Crohn’s disease (183)</td>
</tr>
<tr>
<td>23</td>
<td>Similar dynamics of circulating plasminogen activator inhibitor (PAI)-1 in subjects surviving and dying from sepsis</td>
<td>Demonstrated in CD1 female mice subjected to CLP (185) and post-traumatic sepsis (186); PAI-1 increase correlates with sepsis severity and/or outcome</td>
<td>Demonstrated in surviving and non-surviving patients with sepsis, severe sepsis and/or septic shock (187)</td>
</tr>
<tr>
<td>24</td>
<td>Reduced antioxidant defense and increased oxidative stress in spinal cord injury</td>
<td>Demonstrated in CD1 male mice with spinal cord injury; increase in protein nitration and membrane lipid peroxidation and secondary damage (188,189).</td>
<td>Demonstrated in patients with spinal cord injury; strong and long-term reduction of circulating antioxidants and increase of oxidative stress (190)</td>
</tr>
<tr>
<td>#</td>
<td>Translational phenomenon/response</td>
<td>Specific comments - mouse</td>
<td>Specific comments - human</td>
</tr>
<tr>
<td>----</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>25</td>
<td>Heat shock protein (HSP)-72 serves as a biomarker for early detection of acute kidney injury</td>
<td>Demonstrated in B6/129-j/F2 male iNOS knockout mice subjected to renal injury by ischemia/reperfusion. Increase of renal HSP-72 mRNA and protein correlated with the extent of renal injury (191)</td>
<td>Urinary HSP-72 was significantly increased in patients with clinical acute kidney injury prior to elevation of serum creatinine (192)</td>
</tr>
<tr>
<td>26</td>
<td>Plasma gelsolin is a potential prognostic biomarker for critically ill surgical patients</td>
<td>Exogenous gelsolin infusion reduces brain inflammation and apoptotic signaling and improves survival of mice following major burn injury (193) as well as attenuating burn-induced pulmonary microvascular dysfunction (194)</td>
<td>Low plasma gelsolin levels were associated with increased risk of death occurring in the ICU (195) and correlated with development of multiple organ dysfunction syndrome and fatal outcome in burn patients (196)</td>
</tr>
</tbody>
</table>

*Abbreviations that are not explained in the table: TNF – tumor necrosis factor; CLP – cecal ligation and puncture; NF-κB – nuclear factor kappa B; MARS – mixed anti-inflammatory response syndrome; IL – interleukin; MLCK – myosin L chain kinase; NOD – nucleotide oligomerization domain; LPS – lipopolysaccharide; EDL – endothelial cell-derived lipase; TF – tissue factor; PGC-1α – peroxisome proliferator-activated receptor γ coactivator 1-α; MODS – multi organ dysfunction syndrome; iNOS – inducible nitric oxide synthase; mRNA – messenger ribonucleic acid; ICU – intensive care unit

*www.clinicaltrials.gov* identifier (NCT number)