Quality of Methods Reporting in Animal Models of Colitis

Michael Bramhall, MSc,* Oscar Flórez-Vargas, MSc,* Robert Stevens, PhD,* Andy Brass, PhD,* and Sheena Cruickshank, PhD†

Background: Current understanding of the onset of inflammatory bowel diseases relies heavily on data derived from animal models of colitis. However, the omission of information concerning the method used makes the interpretation of studies difficult or impossible. We assessed the current quality of methods reporting in 4 animal models of colitis that are used to inform clinical research into inflammatory bowel disease: dextran sulfate sodium, interleukin-10−/−, CD45RBhigh T cell transfer, and 2,4,6-trinitrobenzene sulfonic acid (TNBS).

Methods: We performed a systematic review based on PRISMA guidelines, using a PubMed search (2000–2014) to obtain publications that used a microarray to describe gene expression in colitic tissue. Methods reporting quality was scored against a checklist of essential and desirable criteria.

Results: Fifty-eight articles were identified and included in this review (29 dextran sulfate sodium, 15 interleukin-10−/−, 5 T cell transfer, and 16 TNBS; some articles use more than 1 colitis model). A mean of 81.7% (SD = ±7.038) of criteria were reported across all models. Only 1 of the 58 articles reported all essential criteria on our checklist. Animal age, gender, housing conditions, and mortality/morbidity were all poorly reported.

Conclusions: Failure to include all essential criteria is a cause for concern; this failure can have large impact on the quality and replicability of published colitis experiments. We recommend adoption of our checklist as a requirement for publication to improve the quality, comparability, and standardization of colitis studies and will make interpretation and translation of data to human disease more reliable.

(Key Words: IBD, colitis, methods, animal models, checklist)
experiments using animal models. These interventions have largely been successful in raising awareness of flawed methods reporting within the scientific literature, gaining the support of journals, publishing houses, and members of the scientific community. In several cases, publishers have implemented stricter guidelines for methods quality, introduced broad checklists, and removed limitations on word counts for methods reporting. However, there is still a lag between implementation of these measures and adherence to them.

We recently examined the quality of methods reporting in parasitology experiments, highlighting the need for domain-specific guidelines: bespoke checklists tailored by experts that can be used to assess and improve the methods reporting quality within their community. These checklists can be implemented before the point of publication, acting as a barrier to prevent incomplete methods from entering the literature, and also as a review tool for nonexperts when assessing article quality postpublication. Animal models of colitis are numerous, with at least 60 established IBD models currently being used. These models use diverse methods, and the exact mechanics of colitis induction (and the IBD they best model) are poorly understood in some cases. In this article, we aim to briefly summarize the types of colitis model that IBD researchers have at their disposal, highlight some of the problems that experimenters face in producing reliable and robust data from these models, and assess the current quality of methods reporting in published experiments in a subset of available colitis models; scoring them against a checklist of essential and desirable reported methods criteria. The selected criteria cover key aspects that can affect the outcome of colitis in animal models.

We have included checklist criteria relating to 3 broad areas. First, animal sex, age, origin, and housing is considered, which can affect the severity of inflammation, the balance of microbiota in the gut (e.g., strain, diet, acclimation), and animal stress levels (e.g., temperature, animals per cage), and therefore, collectively modulate the severity of induced colitis. Second, factors pertaining to the colitis model, such as genetic modification of animals, origin of chemicals and dosing should be recorded in order for the experiment to be repeatable under the same conditions. Finally, criteria relating to the measurement of colitis, time course of the experiment, and clinical monitoring of animals during the experiment should be reported as standard to determine the success of colitis induction and provide means by which similarity between experiments can be determined for inclusion into systematic reviews and meta-analyses.

Animal Models of IBD

Animal models of colitis have a number of distinct advantages over clinical data when it comes to determining the cause and prevention of IBD. For example, by controlling the onset of inflammation in the laboratory, the failures of immune tolerance, susceptibility genes, and specific proinflammatory pathways involved in triggering colitis can be identified more easily than in a patient admitted with progressive disease and potential comorbidities. Anticolitic preventative measures may also be tested before symptoms occur in an animal model, an impossible task in current treatment of human IBD, where new patients usually only present once the disease reaches clinical significance. The pathway of inflammation can also be accurately modulated in laboratory models to emulate acute or chronic disease depending on the strain of animal used, the mechanism of induction and the use of intervals between deliveries of proinflammatory stimulus.

Although the range of IBD models is diverse, they can be broadly categorized into 4 groups: chemically induced, biologically induced, genetic (including congenic and genetically modified animals), and cell transfer models. We have chosen a cross-section of colitis models to assess methods reporting quality in this field: dextran sulfate sodium (DSS), IL-10 knockout (IL-10−/−), CD4+ CD45RBhigh T cell transfer, and 2,4,6-trinitrobenzene sulfonic acid (TNBS). In addition to animal housing conditions having an impact on the microbiota composition, which itself has a major impact on colitis models, different colitis models have specific criteria that influence their reproducibility as summarized below.

**DSS-induced Colitis Model**

DSS is one of the most commonly used inducers of colitis in animal models, thanks largely to the ease of use and potentially short turnarounds for obtaining results. DSS is typically administered in the drinking water of mice or rats at a dose dependent on the strain of animal, the severity of inflammation desired, and the length of the experiment. Acute and resolving inflammation usually occurs after a single continuous exposure to DSS in drinking water over a week or less, whereas repeated exposure punctuated with recovery periods results in chronic inflammation. The exact mechanism by which DSS induces colitis is still poorly understood, but its primary mode of action seems to chemically interfere with gut mucosa barrier integrity, allowing luminal antigens access to the lamina propria and the proinflammatory cells within. Other factors that can influence the severity and susceptibility of exposure to DSS are the manufacturer and molecular weight of DSS, the strain of animal used (C3H/HeJ and BALB/c mice show increased susceptibility), gender (males are more susceptible), and whether animals are raised in germ-free or specific pathogen-free environments.

**IL-10−/− Chronic Colitis Model**

IL-10 is an anti-inflammatory cytokine that functions to prevent excessive inflammatory and autoimmune pathology. Genome-wide association studies and clinical observations have identified IL-10 as a susceptibility gene for both Crohn’s disease and ulcerative colitis. By employing a number of genetic mechanisms, IL-10 or its receptor have been knocked out or functionally impaired to create several murine animal systems for the study of inflammation. IL-10−/− mice housed under normal conditions develop chronic inflammation in the gut, but mice will remain healthy when housed under germ-free conditions or with a defined selected microbiota and administration of antibiotics can prevent the onset of colitis in IL-10−/− mice. Consequently, to
standardize microbial influence on triggering colitis in the IL-10−/− model, specific enteric microbes such as *Enterococcus faecalis* or *Helicobacter hepaticus* may be used as an inoculum for mice that have been raised in germ-free housing.

**T Cell Transfer Colitis Model**

The T cell transfer model builds on the understanding that T lymphocytes play a pivotal role in the onset of colitis: mediating between antigen presenting cells and generating targeted immune responses to commensal enteric bacteria. In this model, naïve T cells (CD4+ CD45RBhigh or CD4+ CD62L+) are adoptively transferred from wild-type mice into genetically identical mice lacking T cells and B cells (e.g., SCID or RAG−/− mice). The onset of symptoms occurs 2 weeks after T cell transfer in the recipient mice, with pancolitis present from 4 weeks. Due to the extraction, isolation, purification, and injection of adoptive T cells, this model requires a much more complex and labor-intensive protocol than many other IBD models. Factors that influence the resulting colitis include the strain of animal used, the number and viability of T cells transferred, and the presence of B cells in the recipient animals.

**TNBS-induced Colitis Model**

TNBS is a chemical administered rectally in the form of an enema to mice or rats. TNBS is administered in combination with ethanol, which disrupts the mucous barrier, and it is generally thought that TNBS induces colitis by haptenating proteins within the gut, causing them to become preferential targets for immune cells. As with other chemically induced colitis models, the severity of TNBS-induced colitis depends largely on the dosage applied and the strain of animal used.

**Scope of this Study**

A vast amount of clinical and experimental IBD data are available for access: a PubMed search for the Medical Subject Headings (MeSH) term “inflammatory bowel diseases”[MeSH] from the year 2000 to present returns 30,931 articles. Researchers and health professionals cannot possibly hope to consult all the data to make decisions, so we are becoming increasingly reliant on meta-analyses and combinatorial repositories to inform translation from animal experiments to clinical practice: it is vitally important that these processes are built on reliable foundations. This leads us to a pressing need to annotate and accurately record experiments from disparate sources, and this information is often lacking—not only does this prevent construction of well-founded knowledge-base systems, but it also prevents others from fully understanding the validity of results in the context of the experimental setting. How can a reader know whether 2 experiments are comparable if the methods from each experiment are not explicitly clear? In addition, geographical and language barriers or the use of nomenclature experts may prevent the fluid exchange of tacit knowledge, resulting in subtle, yet important, omissions when describing experiments.

To determine whether experiments in the field of primary colitis research are reported with adequate clarity and detail for replication, reproduction, and comparison, we defined a checklist of essential parameters that must be included and desirable parameters that ought to be included when describing experimental animal colitis. We then conducted a PubMed search to obtain a corpus of articles using DSS, IL-10−/−, T cell transfer, or TNBS colitis models for assessing with the checklist. To gather a manageable number of results, we limited the search to studies published after 2000 that conducted a microarray on colonic tissues.

**MATERIALS AND METHODS**

A systematic search was performed following the recommendations of the PRISMA guidelines. Relevant search terms were selected to identify published articles that used 1 (or more) of 4 animal models of colitis: DSS, IL-10−/−, T cell transfer, or TNBS. The search was narrowed down to select only those articles that conducted a microarray on colonic tissues. Assessed criteria were divided into 3 sections in a protocol: aspects relating to the animal and its housing conditions, description of the model of perturbation used and criteria describing the assessment of colitis and the experimental design. The protocol used here for assessing criteria has not been previously published.

The literature search was conducted using PubMed in June 2014 and included articles published in English from January 1, 2000 to June 1, 2014. The search terms included MeSH (Medical Subject Headings) terms and text strings, as outlined in Table 1. The year 2000 was selected as the cutoff due to the emergence of high-throughput analytical techniques becoming more commonplace after the publication of the first draft of the human genome. The DSS model was chosen as this is the most commonly used colitis model. We also selected TNBS as a comparative chemical inducer of colitis, IL-10−/− to represent genetically modified colitis models, and T cell transfer as an example of a model that requires additional, more complex steps in its methods. Biologically induced colitis models, where bacterial or helminthic challenge is used to induce colitis, were not specifically included in this study. However, a number of IL-10−/− articles did include bacterial induction, where a specific cocktail of common murine bacterial strains were used to inoculate germ-free IL-10−/− mice (the checklist is capable of handling biologically induced colitis models). In addition, *Trichuris muris*–induced colitis, while not universally accepted as an IBD model, bears many phenotypic and transcriptional similarities to more traditional IBD models.

However, we chose not to include the *T. muris* infection model in this review as it was covered to some degree in our previous methods quality article.

**Inclusion Criteria**

Primary research articles published in English, within the date constraints, that were returned in the PubMed search were considered for inclusion based on the title and abstract. Reviews, meta-analyses, and experiments that did not use any of the 4 chosen models were excluded. In addition, articles that conducted microarrays on human tissue or primary cell culture tissue only
Animal Models of Colitis

TABLE 1. PubMed Search Terms Used for Each Colitis Model Included in the Systematic Review

| Model          | Search Terms                                                                 |
|----------------|-----------------------------------------------------------------------------|
| DSS            | (Microarray[tw] OR “Microarray Analysis”[Mesh]) AND (“Dextran Sulfate”[Mesh] Dextran sulphate sodium [tw] OR Dextran sulfate sodium [tw] OR DSS [tw]) AND (Inflammatory Bowel Disease*[tw] OR IBD [tw] OR Crohn*[tw] OR Ulcerative Colitis [tw] OR Colitis* [tw] OR Intestin* inflamm* [tw] OR Disease model*[tw] OR “Inflammatory Bowel Diseases”[MeSH] OR “Colitis, Ulcerative”[Mesh] OR “Colitis”[MeSH] OR “Inflammation”[MeSH] OR “Disease Models, Animal”[Mesh]) |
| IL-10^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^-^...
| Group Subgroup                              | No. | Item                                                                 | Essential | Importance | Score | Weight, % |
|--------------------------------------------|-----|----------------------------------------------------------------------|-----------|------------|-------|-----------|
| Information about the animal               |     | Is the species of animal identified? (e.g., mouse)                    | Yes       | High       | 10    | 2.7       |
| Information about the animal               | 1.1 | Is the strain of animal identified? (e.g., C57BL/6)                   | Yes       | High       | 10    | 2.7       |
| Information about the animal               | 1.2 | Is the age of the animal described? (e.g., 12 wks old)                | Yes       | High       | 10    | 2.7       |
| Information about the animal               | 1.3 | Is the gender of the animal described? (e.g., male)                   | Yes       | High       | 10    | 2.7       |
| Information about the animal               | 1.4 | Is the source of animals defined? (e.g., name of supplier or bred in facility) | Yes | High | 10 | 2.7 |
| Information about the animal               | 2.1 | Were animals acclimated to local microbiota? (e.g., housed in identical conditions at least 7 d before experiment start) | Yes | High | 10 | 2.7 |
| Animal housing conditions                  | 3.1 | Is the light/dark cycle described? (e.g., 12 hours light/dark)         | No        | High       | 7     | 1.89      |
| Animal housing conditions                  | 3.2 | Is the temperature described? (e.g., 25°C)                            | No        | Low        | 5     | 1.35      |
| Animal housing conditions                  | 3.3 | Is the humidity described? (e.g., 40%–45%)                            | No        | Low        | 5     | 1.35      |
| Animal housing conditions                  | 3.4 | Is the food/water described? (e.g., regular chow)                     | Yes       | Medium     | 9     | 2.43      |
| Animal housing conditions                  | 3.5 | Is the number of animals per cage described? (e.g., 3 mice per cage)   | No        | Low        | 5     | 1.35      |
| Information about the colitis model        |     | Is the genetic modification identified? (e.g., IL-10<sup>−/−</sup>)    | Yes       | High       | 10    | 2.7       |
| Information about the colitis model        | 4.1 | Is the background strain of the animal described? (e.g., BALB/c)      | Yes       | High       | 10    | 2.7       |
| Information about the colitis model        | 4.2 | Is the chemical used to induce colitis specified? (e.g., DSS)         | Yes       | High       | 10    | 2.7       |
| Information about the colitis model        | 5.1 | Is the molecular weight of the chemical specified? (e.g., 36–50 kDa (DSS only)) | Yes | High | 10 | 2.7 |
| Information about the colitis model        | 5.2 | Is the supplier of the chemical identified? (e.g., Sigma Aldrich)     | Yes       | Low        | 8     | 2.16      |
| Information about the colitis model        | 5.3 | Is the method of induction described? (e.g., dissolved in drinking water) | No | High | 7 | 1.89 |
| Information about the colitis model        | 5.4 | Is the dosage used described? (e.g., 2% wt/vol)                        | Yes       | High       | 10    | 2.7       |
| Information about the colitis model        | 5.5 | Is the medium of inoculation described? (e.g., TNBS in ethanol)        | Yes       | Medium     | 9     | 2.43      |
| Biologically induced colitis model (e.g., bacterial infection) | 6.1 | Is the species of organism identified? (e.g., Helicobacter pylori)  | Yes       | High       | 10    | 2.7       |
| Biologically induced colitis model (e.g., bacterial infection) | 6.2 | Is the strain of organism identified? (e.g., PMSS1)                   | Yes       | High       | 10    | 2.7       |
| Biologically induced colitis model (e.g., bacterial infection) | 6.3 | Are the culture conditions described? (e.g., animal passage or cell culture) | Yes | High | 10 | 2.7 |
| Biologically induced colitis model (e.g., bacterial infection) | 6.4 | Is parasitemia/colonization adequately assessed? (e.g., colon homogenized and plated for colony counting) | Yes | High | 10 | 2.7 |
| Biologically induced colitis model (e.g., bacterial infection) | 6.5 | Is the method of inoculation described? (e.g., oral gavage)            | Yes       | High       | 10    | 2.7       |
| Biologically induced colitis model (e.g., bacterial infection) | 6.6 | Is the dosage used described? (e.g., 10<sup>8</sup> cells)             | Yes       | High       | 10    | 2.7       |
| Adoptive transfer colitis model (e.g., T cell transfer) | 7.1 | Is the cell type being transferred described? (e.g., CD<sup>+</sup> CD45RB<sup>high</sup>) | Yes | High | 10 | 2.7 |
| Adoptive transfer colitis model (e.g., T cell transfer) | 7.2 | Is the species of the donor animal identified? (e.g., mouse)           | Yes       | High       | 10    | 2.7       |
RESULTS

Search Strategy

A total of 58 unique studies were identified for inclusion in the review (see Fig., Supplemental Digital Content 1, http://links.lww.com/IBD/A789). Six of the included articles were applicable to more than 1 of the colitis models and were subsequently included in the datasets for every relevant model (29 DSS, 28–56 15 IL-10−/−, 36,49,50,57–68 5 T cell transfer, 56,69–72 and 16 TNBS35,56,61,73–85; for details of all included studies see Table, Supplemental Digital Content 2, http://links.lww.com/IBD/A790). Duplicate articles were only included once in summary analyses where data from all models are combined. The PubMed searches returned 256 unique articles (54 DSS, 146 IL-10−/−, 42 T cell transfer, and 21 TNBS), 188 of which were rejected based on the title and abstract. A further 10 articles were excluded after assessing the full text of the article, leaving a corpus of 58 articles for analysis.

Quality of Methods Reporting

Each article was assessed for inclusion of the criteria outlined in the quality checklist, which was subdivided into 3 domains: animal, model, and experiment—correlating with subject, perturbation and outcome. The mean weighted score across all colitis models was 81.7% (SD = ±7.038) of criteria reported. By model, articles using the DSS model had the highest quality of methods reporting (mean = 83.30%, SD = ±7.019), and the lowest quality was observed in articles using the T cell transfer model (mean = 73.19%, SD = ±5.328): significantly lower than DSS (P ≤ 0.01) and IL-10−/− (P ≤ 0.05) colitis models (Fig. 1A). Individually, the article with the lowest mean score was 64.05% (T cell transfer model72), and the highest recorded was 94.86% (DSS model52).

For the 4 subsections within “Information about the colitis model,” only the relevant subsections were required. Weights are determined by points attributed to whether the criterion is deemed essential (Yes = 5 or No = 2) plus the level of importance (High = 5, Medium = 4, or Low = 3). The weight for each criterion is then calculated as the percentage of the sum of all scores.

### TABLE 2 (Continued)

| Group                     | Subgroup                          | No. | Item                                                                 | Essential | Importance | Score | Weight, % |
|---------------------------|-----------------------------------|-----|----------------------------------------------------------------------|-----------|------------|-------|-----------|
| Information about the experimental design | Experiment design | 8.1 | Is the time course of the experiment described? (e.g., mice killed after 7 d exposure to DSS) | Yes    | High      | 10    | 2.7     |
|                           | Colitis monitoring and scoring    | 8.2 | Is the method of euthanasia described? (e.g., cervical dislocation) | No       | Medium    | 6     | 1.62    |
|                           |                                   | 8.3 | Is animal weight loss reported? (e.g., as daily % of starting weight) | Yes      | High      | 10    | 2.7     |
|                           |                                   | 8.4 | Is mortality reported? (e.g., survival curve)                        | Yes      | High      | 10    | 2.7     |
|                           |                                   | 8.5 | Is colitis monitored clinically? (e.g., disease activity index)      | No       | High      | 7     | 1.89    |
|                           |                                   | 9.1 | Is colitis scored histologically? (e.g., H&E stain)                  | Yes      | High      | 10    | 2.7     |
|                           |                                   | 9.2 | Is microbiota diversity/population assessed? (e.g., 16S rRNA sequencing) | No    | High      | 7     | 1.89    |
|                           |                                   | 9.3 | Is colon length or weight measured after being killed?              | Yes      | Medium    | 9     | 2.43    |
|                           |                                   | 9.4 | Is the section of gut for analysis identified? (e.g., proximal colon) | Yes      | High      | 10    | 2.7     |

www.ibdjournal.org | 1253

Copyright © 2014 Crohn’s & Colitis Foundation of America, Inc. Unauthorized reproduction of this article is prohibited.
No article reported 100% of all of the criteria on our checklist but 1 article (DSS model) of all the 58 articles assessed successfully reported all essential criteria for every domain.

The best reported domain was the model itself (mean = 95.80%, SD = ±3.018), followed by animal criteria (mean = 64.05%, SD = ±6.992) and experiment criteria (mean = 56.44%, SD = ±10.225). Looking at scores per domain by colitis model, IL-10<sup>−/−</sup> had the highest quality for the animal domain (mean = 70.99%, SD = ±20.194), TNBS had the highest quality for the model domain (mean = 98.94%, SD = ±1.914), and DSS had the highest quality for the experiment domain (mean = 65.78%, SD = ±13.810). The T cell transfer model had the lowest mean scores for all 3 domains (animal = 54.95%, SD = ±7.770; model = 92.00%, SD = ±2.937; experiment = 46.58%, SD = ±14.908) (Fig. 1B). For full details of methods reporting quality for each included study see Tables, Supplemental Digital Content 3-14, http://links.lww.com/IBD/A935, http://links.lww.com/IBD/A936, http://links.lww.com/IBD/A937, http://links.lww.com/IBD/A938, http://links.lww.com/IBD/A939, http://links.lww.com/IBD/A940, http://links.lww.com/IBD/A941, http://links.lww.com/IBD/A942, http://links.lww.com/IBD/A943, http://links.lww.com/IBD/A944, http://links.lww.com/IBD/A945, and http://links.lww.com/IBD/A946.

**DSS-induced Colitis Model**

For DSS colitis, the most poorly reported criteria for the animal domain were food/water, acclimation, animal gender, and animal age (44.83%, 41.38%, 31.03%, and 20.69% of articles failed to report the criteria, respectively). When describing the DSS model itself, 9 articles (31.03%) failed to provide any information about the molecular weight of the DSS used, and 17.24% of articles did not provide information about the supplier of the DSS chemical (Fig. 2). A more detailed examination of the reporting of molecular weight of DSS revealed that of the 20 articles (68.97%) that proved information about the molecular weight of DSS, only 5 (17.24%) used the correct units of measurement: of the remaining 15 articles, 13 (44.83%) provided no units and 2 (6.90%) used incorrect units. Of the 29 articles that used DSS colitis, 24 (82.76%) failed to correctly report the nature of the DSS molecule that they used to induce colitis. The worst reported essential criteria in the experiment design domain were mortality reporting, colon length/weight measurements, animal weight loss, and colitis scoring by histology (72.41%, 51.72%, 20.69%, and 10.34% of articles failed to report these criteria, respectively).

**IL-10<sup>−/−</sup> Chronic Colitis Model**

In the animal domain, the criteria most poorly reported in the articles using the IL-10<sup>−/−</sup> model were very similar to those missing in the DSS model: acclimation, gender, and food/water were the most commonly absent essential criteria (46.67%, 40%, and 33.33% of articles failed to report, respectively). For the IL-10<sup>−/−</sup> model itself, measurement of bacterial colonization in the gut was poorly reported when specific bacterial inoculation was used to induce...
colitis (53.33% failed to report criteria). In addition, 26.67% of IL-10−/− articles did not specify the strain(s) of bacteria used to induce colitis. The worst reported criteria relating to the experimental design were mortality reporting and colon weight/length measurements, which were both absent in 66.67% of articles.

**T Cell Transfer Colitis Model**

For articles using the T cell transfer model, the worst reported criteria in the animal domain were food/water and acclimation (100% and 80% of articles failed to report these criteria, respectively). Gender of animals used was also not specified in 1 of the 5 T cell transfer articles (20%). When describing the T cell transfer model itself, none of the 5 articles described how viability of T cells transferred was measured or whether it was measured at all. For the experimental design, no article using T cell transfer reported mortality of animals used, 60% of articles failed to report colon length/weight measurements, and 40% of articles failed to report animal weight during the experiment.

**TNBS-induced Colitis Model**

Articles using TNBS to induce colitis were the worst for reporting whether animals had been acclimated (87.5% of articles failed to report this criterion). Also, food/water supply and age of animals used was missing in 50% and 25% of articles, respectively. The TNBS model itself was well reported, although 18.75% of articles failed to report the supplier of the TNBS. Similar to the other colitis model, the worst reported essential criteria in the experiment design domain for TNBS were mortality reporting, colon length/weight measurements, animal weight loss, and colitis scoring by histology (75%, 75%, 43.75%, and 37.5% of articles failed to report these criteria, respectively).

**More Recent Articles Have Higher Methods Reporting Quality**

Overall scores have significantly improved year on year \((P = 0.037, r^2 = 0.075)\). T cell transfer is the only model to have a drop in methods reporting quality over time, but this is not significant. DSS and IL-10−/− show a trend toward improved methods reporting quality over time and TNBS overall reporting quality has significantly improved with time \((P = 0.0036, r^2 = 0.4659)\) (Fig. 3A). The improvement in TNBS reporting quality over time has largely come from a significant improvement in the experiment domain \((P = 0.0203, r^2 = 0.3285)\) (Fig. 3B).

**Journal IF Has No Relation to Methods Reporting Quality**

IF was not observed to have a significant impact on methods reporting quality in animal models of colitis (Fig. 3C). When broken down into domains, there was a slight negative

![Figure 3A](https://example.com/fig3a.png)  ![Figure 3B](https://example.com/fig3b.png)  ![Figure 3C](https://example.com/fig3c.png)  ![Figure 3D](https://example.com/fig3d.png)

**FIGURE 3.** A, A significant positive correlation \((P \leq 0.01, r^2 = 0.47)\) is seen between overall methods reporting quality score (%) and year of publication in studies using TNBS-induced colitis. B, The source of this correlation comes largely from the strong positive correlation \((P \leq 0.05, r^2 = 0.33)\) between reporting quality (%) and year of publication within the experimental design subsection in TNBS colitis papers \((n = 16)\). C, IF of the journal of publication had no impact on the overall quality of methods reporting. D, By subdomain, a nonsignificant negative correlation between reduced methods reporting quality and increased IF was observed in the animal domain \((P = 0.0536, r^2 = 0.07)\) \((n = 58)\). Analyses by linear correlation.
Another key factor in determining microbial consistency is diet, with various dietary factors influencing the growth of different bacterial populations in the gut. Again, over half of the studies (53.45%) in our analysis failed to define the chow fed to experimental animals, a factor that can have significant effects on the severity of induced colitis and the microbiota present in the gut. Better standardizations are required for studies where gut microbiota can influence results, and colonization of laboratory animals lasting several weeks, ultimately influencing immune responses in experimental conditions. Movement of animals should be kept to a minimum and laboratory animals require up to 7 days for changes in immune and endocrine parameters to return to baseline before experimental procedures begin, needless to say, these details should be declared in the methods of the study write-up.

FIGURE 4. Bland–Altman plot to assess agreement between 2 experimenters in scoring articles with the minimum information checklist (n = 33). Articles were scored by the second marker, representing at least half the articles assessed for each model. Difference in scores is not significantly different from zero (P = 0.149, r^2 = 0.066).

**VERIFICATION OF CONSISTENCY IN SCORING OF STUDIES**

The second examiner scored 33 of the 58 articles included in the review (DSS = 14, IL-10^-/- = 8, T cell transfer = 3, and TNBS = 8). Differences in scores for the 2 examiners were assessed through a Bland–Altman plot (Fig. 4). Difference in scores between examiners did not differ significantly from zero (P = 0.149, r^2 = 0.066) suggesting that there was no bias in scoring, and articles were scored consistently with the minimum information checklist.

**DISCUSSION**

Chronic inflammation is a complex and poorly understood pathway with important clinical significance both in terms of quality of life and financial impact. It is vitally important that the animal experiments that inform almost all clinical practice are conducted rigorously and published in enough detail for others to benefit from and build upon, which would be in agreement with the principles stated in the 3 Rs (replace, reduce, and refine). To examine the quality of methods reporting in animal models of colitis and determine the potential impact on reliability, replicability, and comparability of studies in this field, we have assessed 4 commonly used animal models of colitis: DSS, IL-10^-/-, T cell transfer, and TNBS. Our results indicate that although these models score well against a checklist of essential criteria, there are still a variety of fundamental criteria that are repeatedly omitted. It is also encouraging to see an improvement over time, even if this effect is quite small. However, the fact that only 1 article from a corpus of 58 reported all essential criteria is a huge cause for concern, 98.3% of articles included in this analysis failed to include sufficient information to accurately repeat the experiment.

In the United Kingdom, death as an endpoint in animal experiments is to be avoided wherever possible. However, mortality and morbidity does occur from time to time and for a variety of reasons, and this should be reported as it will have a significant impact on the data produced and the results of statistical analyses. A statement referring to animal mortality, even if no animal died during the experiment, was one of the worst reported essential criteria from the checklist across all 4 colitis models included in this analysis (48 of 58 articles, 82.76%, failed to include this criterion). Most animal models of colitis are not expected to cause significant morbidity or death, but the lack of reporting, even to confirm that no unexpected deaths occurred, is problematic. When results from animal experiments fail to disclose mortality, bias may be introduced, giving an overly optimistic estimate of the efficacy of the intervention. For example, without adverse event reporting being enforced, there is no obligation for researchers to declare mice that died during an animal study, but failing to declare this information potentially puts the safety of animals and people in future trials at risk. We are not suggesting that the studies included in this review are deliberately obscuring potentially harmful results, and we assume a lack of adverse event reporting reflects an absence of adverse events to report. However, without such a declaration, we cannot say for certain either way. Consequently, animal experiments should align more closely with clinical practice in this regard and declare adverse reactions as a matter of course.

The key role of gut microbiota in the onset and severity of chronic colitis is well defined. Thus, it was surprising that more than half of the studies (63.79%) failed to describe how animals had been acclimated to ensure potential differences in microbiota had been accounted for and controlled. In addition, very few articles specified the use of littermate controls, which would be the ideal gold-standard for controlling baseline equivalence in microbiota populations. It is insufficient to assume animals obtained from the same supplier or reared within the same experimental facility will harbor equivalent microbial populations, as differences can and do exist even within rooms or across facilities. Simple tools to characterize microbiota are available, and, ideally, these should be used to improve standardization and tighten controls within experiments. Alternatively, cohousing or litter mate controls reduce the likely impact of the environment. Additionally, acclimation serves to compensate for stresses involved in transporting animals. Moving cages to a new location in the same facility can have stressful effects on animals lasting several weeks, ultimately influencing immune responses in experimental conditions. Movement of animals should be kept to a minimum and laboratory animals require up to 7 days for changes in immune and endocrine parameters to return to baseline before experimental procedures begin, needless to say, these details should be declared in the methods of the study write-up.
Many studies are conducted where animal facilities are kept at arbitrary numbers with no denomination specific to the experimenters but not necessarily the animals that they house: wild mice spend daytime inactive, nesting at 30 to 32°C and are therefore experiencing cold stress in the majority of animal facilities. Also, in addition to behavioral and immunological changes, mice housed alone will have to endure cooler conditions that mice housed in groups. Severity of colitis in the DSS model is strongly linked to the strain of animal used and the specific conditions that mice housed in groups. Temperature in particular can affect the immune system of mice, with low temperatures triggering immunosuppressive responses.

Many studies are reported the molecular weight of DSS from the same supplier or at the same molecular weight. That only 5 of the 29 DSS articles accurately reported the molecular weight of DSS with the appropriate units is problematic. The presence of arbitrary numbers with no denomination specified or with clearly incorrect units resulting in claims of molecular weight out by orders of magnitude (e.g., kDa instead of Da, or vice versa) in published studies is poor. The increased number of interdisciplinary, non-domain specialists involved in curating and annotating datasets for inclusion in meta-analyses means that this sort of information must be included within the methods of published articles. Authors of studies cannot assume that everyone accessing their study has the expertise to be able to infer the fine details of the protocols they used. Thus, these sorts of errors appearing in the literature suggest potential shortcomings in submission, peer review, and journal editing processes. It is often the responsibility of submitting authors to ensure that there are no errors in a submitted manuscript but peer reviewers ought to be spotting these errors before an article gets to print.

We recommend the continued uptake of methods quality checklists to assist authors and publishers with inclusion of all the relevant methods details that are required to fully interpret data and integrate results into larger analyses. We have provided a domain-specific checklist that can be used in the assessment of methods reporting in any colitis model, and we think this will aid translation of discoveries in animal models into human studies. However, we are aware that by including only microarray studies, we are focusing on a subset of published colitis research. Methods reporting quality for animal models of colitis in general may not reflect the results we have reported here. Also, we have not attempted to address the diversity of experimental design within models or the choice of statistical tests and power calculations used in analysis of data in this field, both of which will impact the feasibility of comparing data from colitis models. It is worth noting that, although all the studies in this review detailed the numbers of mice used per group, none of the studies included any statistical measure of power to justify the number of animals used. This is of concern, as power calculations are important for assessing the validity of statistical tests applied to the data generated and to limit unnecessary use of animals in research.

In conclusion, we have demonstrated that the quality of methods reporting in modeling colitis, while generally appearing high, has serious flaws with long-ranging impact on the translation of primary research into clinical research of IBD. Automated methods, such as computerized histology scoring, may become more commonplace in future, assisting experimenters in standardizing their methods, but more needs to be done to promote and enforce existing guidelines. Animal experimenters have an onus to follow the 3 Rs (replace, reduce, and refine), and better reporting of studies will add value to experimental data produced by animal studies. Implementation of our colitis methods checklist would improve the quality of publications in this field, ensuring animal models, and the data they produce are used effectively to fulfill their maximum usefulness. The pipeline from basic science to clinical practice is filled with examples where success in the laboratory fails to translate into human subjects and improving methods reporting would be an excellent starting point in rectifying this problem at very little cost or effort.

ACKNOWLEDGMENTS

We wish to thank Cath Booth and James Wilson at Epistem, Dr. Joanne Pennock and Professor Werner Muller at the University of Manchester for support in this study.

REFERENCES

1. Van Limbergen J, Radford-Smith G, Satsangi J. Advances in IBD genetics. Nat Rev Gastroenterol Hepatol. 2014;11:372–385.
2. Mizoguchi A. Animal models of inflammatory bowel disease. In: Conn PM, ed. Progress in Molecular Biology and Translational Science. London, United kingdom: Academic Press; 2012:105:263–320.
3. Goyal N, Rana A, Aklawat A, et al. Animal models of inflammatory bowel disease: a review. Inflammopharmacology. 2014;22:219–233.
4. Kamada N, Sato S, Chen GY, et al. Role of the gut microbiota in immunity and inflammatory disease. Nat Rev Immunol. 2013;13:321–335.
5. Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. Lancet. 2014;383:267–276.
6. Kilkenney C, Browne WJ, Cuthill IC, et al. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. PLoS Biol. 2010;8:e1000412.

7. Noorden RV. Parasite test shows where validation studies can go wrong [Nature News web site]. 2014. Available at: http://www.nature.com/news/parasite-test-shows-where-validation-studies-can-go-wrong-1. 16527/WT_mc_id=TWT_NatureNews. Accessed December 19, 2014.

8. Taylor CF, Field D, Sansone SA, et al. Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. Nat Biotechnol. 2008;26:889–896.

9. Landis SC, Amara SG, Asadullah K, et al. A call for transparent reporting to minimize the predictive value of preclinical research. Nature. 2012; 490:187–191.

10. Raising standards. Nature. 2010;468:1052–1054.

11. Landis SC, Amara SG, Asadullah K, et al. A call for transparent reporting to minimize the predictive value of preclinical research. Nature. 2012; 490:187–191.

12. Emsley P, Cheeseman HM, Pannetier C, et al. The crystal structure of the human angiotensin-converting enzyme: implications for drug design. Nature. 1990;347:545–55 1.

13. Cho JJ, Hwang JK, Chun HS, et al. Melatonin attenuates dextran sodium sulfate induced colitis with sleep deprivation: possible mechanism by microarray analysis. Dig Dis Sci. 2014;59:1134–1141.

14. Coburn LA, Gong X, Singh K, et al. L-arginine supplementation improves responses to injury and inflammation in dextran sulfate sodium-induced colitis. PLoS One. 2012;7:e33546.

15. Reinkvam DH, Derrien M, Islam R, et al. Epithelial-microbial crosstalk and HPA axis responsiveness. J Physiol Gastrointest Liver Physiol. 2013;34:345–364.

16. Hufeldt MR, Nielsen DS, Vogensen FK, et al. Changes in mucosal homeostasis predispose NHE3 knockout mice to increased susceptibility to DSS-induced epithelial injury. Gastroenterology. 2009;137:965–975, 975e1–975e10.

17. Noorden RV. Parasite test shows where validation studies can go wrong [Nature News web site]. 2014. Available at: http://www.nature.com/news/parasite-test-shows-where-validation-studies-can-go-wrong-1. 16527/WT_mc_id=TWT_NatureNews. Accessed December 19, 2014.

18. Emsley P, Cheeseman HM, Pannetier C, et al. The crystal structure of the human angiotensin-converting enzyme: implications for drug design. Nature. 1990;347:545–55 1.

19. Cho JJ, Hwang JK, Chun HS, et al. Xanthorrhizol attenuates dextran sulfate sodium-induced colitis via the modulation of the expression of inflammatory genes in mice. Life Sci. 2011;88:864–870.

20. Chung SH, Park YS, Kim OS, et al. Melatonin attenuates dextran sodium sulfate induced colitis with sleep deprivation: possible mechanism by microarray analysis. Dig Dis Sci. 2014;59:1134–1141.

21. Coburn LA, Gong X, Singh K, et al. L-arginine supplementation improves responses to injury and inflammation in dextran sulfate sodium-induced colitis. PLoS One. 2012;7:e33546.

22. Fang K, Bruce M, Pattullo CB, et al. Temporal genomewide expression profiling of DSS colitis reveals novel inflammatory and angiogenesis genes similar to ulcerative colitis. Physiol Genomics. 2011;43:43–56.

23. Hamilton MJ, Sinnamon MJ, Lyng GD, et al. Essential role for mast cell tryptase in acute experimental colitis. Proc Natl Acad Sci U S A. 2011; 108:290–295.

24. Hansen JJ, Holt L, Sartor RB. Gene expression patterns in experimental colitis in IL-10-deficient mice. Inflamm Bowel Dis. 2009;15:890–899.

25. Hontecillas R, Horne WT, Climenti M, et al. Immunoregulatory mechanisms of macrophage PPAR-gamma in mice with experimental inflammatory bowel disease. Mucosal Immunol. 2011;4:304–313.

26. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. BMJ. 2009;339:b2535.

27. Noorden RV. Parasite test shows where validation studies can go wrong [Nature News web site]. 2014. Available at: http://www.nature.com/news/parasite-test-shows-where-validation-studies-can-go-wrong-1. 16527/WT_mc_id=TWT_NatureNews. Accessed December 19, 2014.

28. Emsley P, Cheeseman HM, Pannetier C, et al. The crystal structure of the human angiotensin-converting enzyme: implications for drug design. Nature. 1990;347:545–55 1.

29. Breynaert C, Dresselaers T, Perrier C, et al. Unique gene expression and inflammatory and reduces DSS-induced colitis in mice. J Physiol Gastrointest Liver Physiol. 2013;34:345–364.

30. Cho JY, Chi SG, Chun HS, et al. Xanthorrhizol attenuates dextran sulfate sodium-induced colitis via the modulation of the expression of inflammatory genes in mice. Life Sci. 2011;88:864–870.

31. Cho JY, Hwang JK, Chun HS, et al. Xanthorrhizol attenuates dextran sulfate sodium-induced colitis with sleep deprivation: possible mechanism by microarray analysis. Dig Dis Sci. 2014;59:1134–1141.
association with mucosal epigenetic and microbiomic changes. Hum Mol Genet. 2011;20:1687–1696.

56. te Velde AA, de Kort F, Sterrenburg E, et al. Comparative analysis of colonic gene expression of three experimental colitis models mimicking inflammatory bowel disease. Inflamm Bowel Dis. 2007;13:325–330.

57. Barnett MP, McNabb WC, Cookson AL, et al. Changes in colon gene expression associated with increased colon inflammation in interleukin-10-deficient mice inoculated with Enterococcus species. BMC Immunol. 2010;11:39.

58. de Buhr MF, Mahler M, Geffers R, et al. Cd14, Gbp1, and Pla2g2a: three major candidate genes for experimental IBD identified by combining QTL and microarray analyses. Physiol Genomics. 2006;25:426–434.

59. Edmunds SJ, Roy NC, Davy M, et al. Effects of kiwifruit extracts on colonic gene and protein expression levels in IL-10-deficient mice. Br J Nutr. 2012;108:113–129.

60. Hemmerling J, Heller K, Hormannsperger G, et al. Fetal exposure to maternal inflammation does not affect postnatal development of genetically-driven ileitis and colitis. PLoS One. 2014;9:e92337.

61. Huang Z, Shi T, Zhou Q, et al. mir-141 regulates colonic leukocytic trafficking by targeting CXCL12-beta during murine colitis and human Crohn’s disease. Gut. 2014;63:1247–1257.

62. Knoch B, Barnett MP, Cooney J, et al. Dietary oleic acid as a control fatty acid for polyunsaturated fatty acid intervention studies: a transcriptionomics and proteomics investigation using interleukin-10 gene-deficient mice. Biotechnol J. 2010;5:1226–1240.

63. Knoch B, Barnett MP, McNabb WC, et al. Dietary arachidonic acid-mediated effects on colon inflammation using transcriptome analysis. Mol Nutr Food Res. 2010;54(suppl 1):S62–S74.

64. Knoch B, Barnett MP, Zhu S, et al. Genome-wide analysis of dietary eicosapentaenoic acid- and oleic acid-induced modulation of colon inflammation in interleukin-10 gene-deficient mice. J Nutrigenet Nutrigenomics. 2009;2:9–28.

65. Kuo SM, Chan WC, Hu Z. Wild-type and IL10-null mice have differential colonic epithelial gene expression responses to dietary supplementation with symbiotic Bifidobacterium animalis subspecies lactis and inulin. J Nutr. 2014;144:245–251.

66. Reiff C, Delday M, Rucklidge G, et al. Balancing in inflammation switches the differentiation of Ly6Chi monocytes from anti-inflammatory dendritic cells in the colon. J Exp Med. 2012;209:1069–1074.

67. Guzman J, Yu JG, Sunthises Z, et al. ADOA3R as a therapeutic target in experimental colitis: proof by validated high-density oligonucleotide microarray analysis. Inflamm Bowel Dis. 2006;12:766–789.

68. Kremmer E, Mariman E, van Erk M, et al. Temporal colonic gene expression profiling in the recurrent colitis model identifies early and chronic inflammatory processes. PLoS One. 2012;7:e50388.

69. Mariman E, Kremmer E, van Erk M, et al. Gene expression profiling identifies mechanisms of protection to recurrent trinitrobenzene sulfonic acid colitis mediated by probiotics. Inflamm Bowel Dis. 2012;18:1424–1433.

70. Martinez-Augustin O, Merlos M, Zarzuazo A, et al. Disturbances in metabolic, transport and structural genes in experimental colonic inflammation in the rat: a longitudinal genomic analysis. BMC Genomics. 2008;9:490.

71. Nur T, Peijnenburg AA, Noteborn HP, et al. DNA microarray technology reveals similar gene expression patterns in rats with vitamin a deficiency and chemically induced colitis. J Nutr. 2002;132:2131–2136.

72. Rivera E, Flores I, Appleyard CB. Molecular profiling of a rat model of colitis: validation of known inflammatory genes and identification of novel disease-associated targets. Inflamm Bowel Dis. 2006;12:950–966.

73. Wu F, Chakravarti S. Differential expression of inflammatory and fibrogenic genes and their regulation by NF-kappaB inhibition in a mouse model of chronic colitis. J Immunol. 2007;179:6988–7000.

74. Yamamoto S, Isuzugawa K, Takahashi Y, et al. Intestinal gene expression profile in TNBS-treated mice using genechip and subtractive cDNA analysis: implications for Crohn’s disease. Biol Pharm Bull. 2005;28:2046–2053.

75. Zhou W, Cao Q, Peng Y, et al. FoxO4 inhibits NF-kappa B and protects mice against colonic injury and inflammation. Gastroenterology. 2009;137:1403–1414.

76. Zwiers A, Fuss JJ, Leijen S, et al. Increased expression of the tight junction molecule claudin-18 A1 in both experimental colitis and ulcerative colitis. Inflamm Bowel Dis. 2008;14:1652–1659.

77. NC3Rs. The 3Rs [NC3Rs web site]. Available at: http://www.nc3rs.org.uk/the-3rs. Accessed 23 October, 2014.

78. Home Office. Guidance on the Operation of the Animals (Scientific Procedures) Act 1986. London, United Kingdom: HMSO; 2014.

79. Clarke M. Standardising outcomes for clinical trials and systematic reviews. Trials. 2007;8:39.

80. Couzin-Frankel J. When mice mislead. Science. 2013;342:922–925.

81. Check Hayden E. Misleading mouse studies waste medical resources. Nature News. 26 March, 2014.

82. Rogers GB, Kozlowska J, Keeble J, et al. Functional divergence in gastrointestinal microbiota in physically-separated genetically identical mice. Sci Rep. 2014;4:5437.

83. Obernberger JA, Baldwin RL. Establishing an appropriate period of acclimatization following transportation of laboratory animals. ILAR J. 2006;47:364–369.

84. Tremaroli V, Backhed F. Functional interactions between the gut microbiota and host metabolism. Nature. 2012;489:242–249.

85. Laying better plans for mice. Nat Biotechnol. 2013;31:263.

86. Fish EN. The X-files in immunity: sex-based differences predispose immune responses. Nat Rev Immunol. 2008;8:737–744.

87. Karp CL. Unstresssing intemperate models: how cold stress undermines mouse modeling. J Exp Med. 2012;209:1069–1074.

88. Kokoszka KM, Capellos KM, Derickowski J, et al. Baseline tumor growth and immune control in laboratory mice are significantly influenced by subthermoneutral housing temperature. Proc Natl Acad Sci U S A. 2013;110:20176–20181.

89. Rabin BS, Lyte M, Epstein LH, et al. Alteration of immune competency by number of mice housed per cage. Ann N Y Acad Sci. 1987;496:492–500.

90. Kitajima S, Takuma S, Morimoto M. Histological analysis of murine colitis induced by dextran sulfate sodium of different molecular weights. J Pathol. 2006;209:103–112.

91. Rogers GB, Kozlowska J, Keeble J, et al. Functional divergence in gastrointestinal microbiota in physically-separated genetically identical mice. Sci Rep. 2014;4:5437.