Incidence, severity, mortality and confounding factors for dissecting AAA detection in angiotensin II-infused mice: a meta-analysis

Bram Trachet[1,2], Rodrigo Fraga-Silva[2], Philippe Jacquet[3,4], Nikolaos Stergiopulos[2], Patrick Segers[1]

[1] IBiTech - bioMMeda, Ghent University-iMinds Medical IT, Ghent, Belgium

[2] Institute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

[3] Bioinformatics and Biostatistics Core Facility, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

[4] Swiss Institute of Bioinformatics, Lausanne, Switzerland

Bram Trachet
LHTC STI IBI EPFL
BM 5128 Station 17
CH-1015 Lausanne (Switzerland)
Tel: +41 21 693 83 81
Fax: +41 21 693 96 35
bram.trachet@ugent.be

Incidence, severity, mortality and confounding factors for dissecting AAA detection in angiotensin II-infused mice: a meta-analysis

Bram Trachet; Rodrigo A. Fraga-Silva; Philippe Jacquet; Nikolaos Stergiopulos; Patrick Segers

Abstract

While angiotensin II-infused mice are the most popular model for preclinical aneurysm research representative data on incidence, severity and mortality of dissecting abdominal aortic aneurysms (AAAs) have never been established, and the influence of confounding factors is unknown. In this review, we performed a meta-analysis including 194 manuscripts representing 1679 saline-infused, 4729 non-treated angiotensin II-infused and 4057 treated angiotensin II-infused mice. Incidence (60%) and mortality (20%) rates are reported overall as well as for grade I (22%), grade II (26%), grade III (29%) and grade IV(24%) aneurysms. Dissecting AAA incidence was significantly ($p<0.05$) influenced by sex, age, genetic background, infusion time and dose of angiotensin II. Mortality was influenced by sex, genetic background and dose, but not by age or infusion time. Surprisingly, both incidence and mortality were significantly different ($p<0.05$) when comparing angiotensin II-infused mice in descriptive studies (56% incidence, 19% mortality) with angiotensin II-infused mice that served as control animals in treatment studies designed to either enhance (35% incidence, 13% mortality) or reduce (73% incidence, 25% mortality) dissecting AAA formation. After stratification to account for confounding factors (selection bias) the observed effect was still present for incidence, but not for mortality. Possible underlying causes are detection bias (non-uniform definition for detection and quantification of dissecting AAA in mice) or publication bias (studies with negative results, related to incidence in the control group, not being published). In conclusion, our data indicate that treatment studies using angiotensin II-infused mice should be interpreted with caution and provide a new meta-analysis-based reference for incidence and mortality of dissecting AAA in this mouse model

Introduction

In clinical practice, abdominal aortic aneurysm (AAA) is defined as a focal dilatation of the aortic diameter larger than 1.5 times the original size [1]. Due to the asymptomatic nature of the disease human data are scarce, especially at an early stage. Therefore animal models of aneurysm formation are often used for pre-clinical research [2]. The angiotensin II-infused mouse model is the most popular mouse model for aneurysm research because it reproduces many important features such as macrophage infiltration, medial elastolysis, luminal expansion and thrombus formation[3-5].

Since the first paper published by Daugherty *et al*. in 2000 [3], this mouse model has been the subject of numerous papers, ranging from the genetic [6-10] over the molecular [2, 11-25] to the macroscopic level [26-34]. Lesion severity has been reported to vary substantially within this mouse model, leading to a subdivision into Grade I-Grade IV aneurysms based on morphological characteristics that were visually observed on excised tissue samples [20]. Over time, a wide range of incidence, severity and mortality rates has been reported for this mouse model but reference data are lacking. Moreover, while several confounding factors (sex, age, diet) seem to be the same as in human disease, their effect on murine aneurysm pathology has never been quantified.

The angiotensin II mouse model is particularly popular in so-called treatment studies, in which the effect of pharmaceutical or genetic interventions on aneurysm incidence, aneurysm severity and aneurysm-related mortality are assessed. Since 2001, 143 manuscripts have studied the effect of some kind of treatment in angiotensin II-infused mice, often quantifying aneurysm presence in terms of the luminal or external diameter change. But despite its ubiquitous use, some important differences with human AAAs exist: the latter are located on the infrarenal rather than the suprarenal aspect of the abdominal aorta and are characterized by a progressive evolution rather than repeated interparietal ruptures [5, 35]. In order to avoid confusion, we

therefore refer to the angiotensin II-induced lesions as `dissecting AAAs` rather than AAAs throughout the current manuscript [36].

Here, we present a meta-analysis of all manuscripts (up to January 1st, 2015) in which mice have been infused with angiotensin II in order to provoke dissecting AAA formation. In first instance, our goal was to define reference values for (i) the occurrence of dissecting AAA, (ii) the occurrence of different severity levels within animals developing dissecting AAA (categorized from grade I-grade IV [20]) and (iii) the mortality rate during angiotensin II-infusion, related to aneurysm rupture. We also report the influence of different confounding factors such as sex, age, diet, genetic background and dose of angiotensin II on incidence and mortality rates. The second part of the manuscript focuses on the statistically significant difference that was discovered in the incidence and mortality rates of non-treated mice in treatment studies that were designed to enhance dissecting AAA, on the one hand, and treatment studies that were designed to reduce dissecting AAA, on the other hand. This difference seems to imply that many of the reported incidence and mortality rates in literature were biased by the purpose of the study. We present a stratified analysis to investigate to what extent this bias could be attributed to confounding environmental factors, and to what extent it was related to the variability that exists in the criteria used to define AAA presence. Based on these findings, we propose alternative criteria to quantify dissecting AAA presence, rather than focusing on diameter alone. As such, this manuscripts aims to provide deeper insight into the interpretation of angiotensin II-induced dissecting AAA in the past, and provide improved guidelines for its interpretation in the future.

Methodology

The meta-analysis included all publications available in Pubmed until January 1st, 2015, using the search terms "mice" or "mouse", "aneurysm" and "angiotensin". Exclusion criteria were (i) the use of different mouse models than angiotensin II infusion (e.g. elastase perfusion or genetic modification without angiotensin II infusion), (ii) reporting on locations of aneurysm formation different from the abdominal aorta (e.g. intracranial or ascending aneurysms), (iii) reports focused on human aneurysm or aneurysm-related disorders (e.g. Marfan or Loeys-Dietz syndrome), (iv) review articles, (v) letters to the editor, (vi) manuscripts that re-used mice from a previous manuscript (and had thus already been included in the analysis) and (vii) manuscripts for which no access could be obtained.

From the remaining 194 manuscripts the following data were retrieved from the methods and results sections: sex (male/female), diet (normal/high fat), genetic background (ApoE$^{-/-}$ / LdL$^{-/-}$ / C57Bl\6 wild type, C57Bl\6 wild type with anti-TGF-β antibodies, C57Bl\6 wild type with co-infusion of β-apn), age (ranging from 4 to 72 weeks), angiotensin II infusion dose (ranging from 150 to 3000 ng/min/kg), angiotensin II infusion time (ranging from 3 to 84 days), the location where dissecting AAA diameters were quantified (lumen diameter or external diameter), the threshold of diameter increase used to define dissecting AAA incidence (ranging from 110 to 200%) and the reference to which dissecting AAA diameters were compared in order to define dissecting AAA (suprarenal diameters in saline-infused control animals, baseline suprarenal diameters prior to angiotensin II infusion, diameters of a non-diseased segment proximal to the dissecting AAA, diameters of a non-diseased segment distal to the dissecting AAA, or lumen diameters at locations where the dissecting AAA was quantified at the external diameter). Each study was classified into one of three categories, depending on whether its main goal was to describe aneurysm formation, to enhance aneurysm formation, or to reduce aneurysm formation in angiotensin II-infused mice. In manuscripts designed to

describe murine aneurysm formation, angiotensin II-infused and saline-infused mice were the only study groups. In manuscripts designed to enhance or reduce dissecting AAA formation, the angiotensin II-infused mice usually functioned as controls to a third group, in which the effect of dissecting AAA was reduced or enhanced by pharmacological or genetic treatment. Two studies could not be categorized into any of the three categories. For each study the number of saline-infused, the number of angiotensin-infused and the number of treated mice was retrieved, as well as the reported incidence rate of mice that developed an dissecting AAA and mice that died of aneurysm rupture in each category. If the incidence or rupture rates were reported in the form of bar plots or percentages an estimate was made based on the available data, and if group size was reported as a range the mean value was used. Aneurysm mortality was defined as sudden death during the infusion period, since the pathological confirmation of an internal bleeding due to aneurysm rupture nor its precise location (hemoabdomen versus hemothorax) were mentioned in most manuscripts. Following the most common procedure in literature, incidence rates include mortality rates. In several manuscripts a number of different treatment groups were reported. If any of the confounding factors were different between groups, these were treated as separate studies. If not, incidence rates from different groups were added within the same study.

Statistics

Odds ratios were calculated for each confounding factor with respect to its reference situation. Odds ratios for categorical variables such as sex, diet and genetic background compared dissecting AAA incidence or mortality of mice with the less frequently occurring value of the confounding factor to the incidence of 'reference mice' for that factor (i.e. female was compared to male, high fat diet to normal diet, C57Bl\6 to ApoE$^{-/-}$). Continuous variables such as age, angiotensin II dose and angiotensin II infusion time were first categorized (young: 4-8

weeks, old: 25-84 weeks; low dose: 125-749 ng/kg/min, high dose: 1251-3000 ng/kg/min; low

infusion time: < 7 days) and odds ratios subsequently compared dissecting AAA incidence and

mortality for these less frequent values to their respective reference situation (adult: 9-24 weeks;

normal dose: 750-1250 ng/kg/min; normal infusion time: 7-28 days). Reference measurements

to define dissecting AAA incidence were considered to be external diameters using a threshold

of a 150% increase in comparison with saline-infused controls. Odds ratios were considered

significant (p<0.05) if the 95% confidence interval did not include the value OR=1.

Since we observed an unexpected dependency of both dissecting AAA incidence and mortality

on the design of the study, we subsequently performed a stratified analysis to find to what extent

this dependency was related to the presence of confounding factors. First a Fisher's exact test

was performed to identify the environmental factors that were correlated with study design on

both a study level (i.e. quantifying how the number of studies associated to each study design

was related to each confounding factor) and a mouse level (i.e. quantifying how the

accumulated aneurysm incidence and mortality for each study design was related to each

confounding factor). Then a stepwise stratified analysis was performed in which the effect of

these confounding factors on the odds ratios was investigated at an individual mouse level. For

each factor, only those mice representing the respective reference situation were withheld (i.e.

after stratification for age only adult mice were withheld, after stratification for genetic

background only Apoe$^{-/-}$ mice were withheld, etc.). If any of the confounding factors was not

reported, or if either the incidence rate or the number of animals were missing, the animals were

excluded from the stratified analysis. First we corrected the odds ratios for those factors that

were significantly correlated to study design on both a study and a mouse level. Since a

significant effect was still visible, stratification was subsequently expanded to confounding

factors that were only significantly correlated with study design on a mouse level. Finally

stratification also included measurement related factors.

Results

Metrics of the meta-analysis

In total 291 papers were analyzed. Forty-two manuscripts were excluded since they studied different mouse models, 21 manuscripts focused on human disease, 28 did not report original research and 2 manuscripts re-used mice from a previous manuscript, while for 4 manuscripts no access could be obtained. Thirty-five manuscripts reported several studies with varying incidence rates or confounding factors, each of which was considered as a separate entry for the meta-analysis. In total, this approach led to a total number of 252 studies that were extracted from 194 manuscripts. An overview can be found in table S1 (supplementary data). These manuscripts were published in 64 different journals, with an average journal impact factor of 5.7 ± 3.7. Overall, the meta-analysis represents 1679 saline-infused, 4729 non-treated angiotensin II-infused and 4057 treated angiotensin II-infused mice.

Dissecting AAA incidence, mortality and grade I-IV morphology

Overall, 60% of the animals developed a dissecting AAA while the mortality during angiotensin II infusion (presumably related to transmural rupture) was found to be 20% (Table 1). In 33 studies representing 429 dissecting AAAs, aneurysm shapes were categorized into 4 different groups, ranging from grade I to grade IV, according to the morphology criterion first described by Daugherty et al [20]. In angiotensin II-infused mice that did not receive any additional treatment, we found an overall incidence of 22% grade I aneurysms, 26% grade II aneurysms, 29% grade III aneurysms, and 24% grade IV aneurysms.

The influence of environmental factors on dissecting AAA incidence and mortality

As reported previously, dissecting AAA incidence was significantly lower for females than for males (Table 2, Figure 1). Young mice experienced a significantly lower incidence than adult

mice, while old mice showed a significantly higher incidence than adults. Significantly less dissecting AAAs were observed in angiotensin II-infused wild-type C57Bl\6 mice than in ApoE $^{-/-}$ mice , and also LdL $^{-/-}$ mice had significantly lower incidence than ApoE $^{-/-}$ mice. C57BL\6 mice that were co-infused with angiotensin II and β-apn had an incidence rate that was not different from that of angiotensin II–infused ApoE $^{-/-}$ mice, while C57Bl\6 mice that were infused with angiotensin II and injected with anti-TGF-β antibodies had a significantly higher incidence than angiotensin II-infused ApoE $^{-/-}$ mice.

Dissecting AAA incidence was significantly lower for mice receiving a low dose of angiotensin II (< 750 ng/kg/min) compared to those receiving a normal dose, but the difference in incidence for mice that were infused with higher doses of angiotensin II (> 1250 ng/kg/min) was not statistically significant. A high-fat diet did not result in a statistically significant difference in incidence compared to normal rodent chow.

Rupture rate followed the same trend as incidence for nearly all reported confounding factors. Mortality during angiotensin II infusion was significantly higher in male mice, anti-TGF-β injected mice, and treated mice in studies designed to enhance dissecting AAA. Mortality during angiotensin II-infusion was significantly lower in wild type C57Bl\6 mice, LdL $^{-/-}$ mice, mice receiving a low dose of angiotensin II, and treated mice in a study designed to reduce dissecting AAA. No significant difference in mortality was found in mice co-treated with β-apn, in mice receiving a high dose of angiotensin II or in mice on a high fat diet. In fact, the only confounding factor that had a different influence on mortality than on aneurysm incidence was age. While dissecting AAA incidence was significantly lower in young mice than in adult mice, no significant difference in mortality could be detected. Similarly, no statistically significant difference in mortality was found between old and adult mice.

The enhancement of reductions and the reduction of enhancements: a short description

The most surprising result from our meta-analysis lies in the strong relation that was found between study design and incidence of dissecting AAA in non-treated, angiotensin II-infused mice. In manuscripts that were designed to describe murine aneurysm formation the incidence rate in angiotensin II-infused mice was 56%. However, in manuscripts that were designed to enhance murine aneurysm formation the incidence rate in non-treated, angiotensin II-infused mice (that were used as control animals in these studies) was only 35%. Compared to aneurysm-describing studies, this difference is highly significant and corresponds to an odds ratio of 0.43 (Table 2, Figure 1). On the other hand, in manuscripts that were designed to reduce murine aneurysm formation the incidence rate in non-treated mice (that equally served as control animals in these studies) was as high as 73%. Compared to aneurysm-describing studies, this difference was highly significant as well, and corresponded to an odds ratio of 2.12 (Table 2, Figure 1). When comparing studies designed to enhance dissecting AAA directly to those designed to reduce it, the difference in dissecting AAA incidence of non-treated mice was even more pronounced (OR=0.20).

A similar trend could be observed in dissecting AAA mortality rates. In manuscripts that were designed to describe murine aneurysm formation the mortality was 19%. However, in manuscripts that were designed to enhance murine aneurysm formation the mortality in non-treated mice was only 13%, while it was 25% in manuscripts that were designed to reduce murine aneurysm formation. Both odds ratios were significantly different from 1 (Table 2, Figure 1).

Stratification for environmental factors

A Fisher's exact test indicated that study design significantly correlated with sex, genetic background, and angiotensin II dose ($p<0.05$, Table 3). More specifically, researchers tend to choose more often male mice and avoid low doses of angiotensin II in studies designed to

reduce dissecting AAA (Figure 2). On the other hand researchers conducting studies designed to enhance dissecting AAA more often selected LdL $^{-/-}$ or C57Bl/6 mice and used lower doses of angiotensin II. In first instance we performed a step-wise stratified analysis in which we accounted for the effect of these confounding factors on both incidence and mortality (Table 4, Figure 3).

After stratification for sex, background and dose (the environmental factors significantly correlated with study design on both a study level and a mouse level, see Table 3) incidence was still significantly lower for studies designed to enhance dissecting AAA. Similarly, incidence was still significantly higher for studies designed to reduce it. Further stratification for infusion time, age and diet (all of which were only significantly correlated with study design on a mouse level) did not influence the odds ratios in either case (Figure 3, top). Interestingly, the difference in mortality rates for studies designed to enhance dissecting AAA disappeared after stratification for sex and background. Similarly, the difference in mortality rates for studies designed to reduce dissecting AAA disappeared after stratification for sex, background, dose, infusion time and diet (Figure 3, bottom).

The influence of measurement related factors on dissecting AAA incidence and mortality

We were able to extract the used criterion for aneurysm incidence from 115 out of 194 manuscripts. In the remaining studies, the exact criterion for dissecting AAA was either not mentioned, or we did not manage to locate it within the manuscript. Of the analyzed studies only 84% quantified dissecting AAA dimensions, while 16% determined incidence based on visual inspection of the tissue. Surprisingly, the different criteria used to define incidence criteria result in no less than 31 potential definitions for dissecting AAA incidence (visual inspection plus any permutation of 3 diameter thresholds, 5 reference locations and 2 measurement locations; see Figure 2).

Figure 4 shows the odds ratios of those measurement related factors that deviate from the most frequently occurring 'reference' measurement. Incidence is shown to be significantly lower when the threshold to define dissecting AAA presence is put higher than 150%, and when lesion diameters are compared to a proximal segment of the same aorta, rather than saline-infused controls. The incidence of dissecting AAA is significantly higher when the aneurysm diameter is measured at the luminal border (i.e. in vivo) than when external diameters are measured. Also, a significantly higher incidence was found when comparing the dissecting AAA segment to baseline suprarenal diameters (prior to angiotensin II infusion), and when comparing it to a distal segment. Mortality rates showed to be much less dependent on the measurement related factors.

Stratification for measurement related factors

A Fisher's exact test indicated that study design significantly correlated with the reference to which the suprarenal diameter was compared and also with the percentage increase in diameter that was used as a threshold to define the presence of dissecting AAA (Table 3). More specifically, researchers tend to use less stringent criteria for diameter increase in studies designed to reduce dissecting AAA (Figure 2). Also, researchers conducting treatment studies (be it to enhance or to reduce dissecting AAA) more often measure the outer diameter (rather than the lumen) and compare suprarenal diameter values more often to those of saline-infused mice than researchers in descriptive studies. The difference is most outspoken in studies designed to reduce dissecting AAA (Figure 2).

For studies designed to enhance dissecting AAA, additional stratification for measurement related factors removes the difference in incidence with studies designed to describe dissecting AAA (Table 4, Figure 5, top). This is the case when withholding only those studies in which the suprarenal diameter is compared to the diameter of saline-infused mice, but also when

withholding only studies in which a diameter criterion of 150% was used, and when withholding only studies in which the external diameter was measured (rather than the luminal diameter). Conversely, for studies designed to reduce dissecting AAA this stratification led to a further increase in the odds ratio for both the saline reference and the external diameter. For these studies the difference in incidence only became (borderline) insignificant when only those studies in which a diameter criterion of 150% was used were selected. In order to account for the low amount of mice remaining after such stringent stratification, we subsequently visualized the effect of stratification for all values of each measurement factor, ignoring any effect of environmental confounding factors. For studies designed to reduce dissecting AAA, the odds ratio only decreased for studies that compared dissecting AAA diameters to a baseline reference, for those using a threshold < 150%, or for those measuring the luminal diameter (Figure 5, bottom).

Discussion

The last years a plethora of studies have used angiotensin II-infused mice to study dissecting AAA in mice: 93 of the 194 studies analyzed in this meta-analysis were published between 2012 and 2015. In the first part of this meta-analysis we provide the research community with reference values concerning the incidence and mortality rates of dissecting AAA in this mouse model, and report to what extent these values are influenced by environmental confounding factors such as sex, age, diet, genetic background and angiotensin II dose. In the second part we focus on the apparent discrepancy in reported incidence rates for studies designed to enhance and studies designed to reduce dissecting AAA. We investigate possible causes for this potential bias, and propose alternative criteria for dissecting AAA detection in future research.

Dissecting AAA incidence and mortality and their confounding factors

Our meta-analysis is the first to come up with incidence, severity and mortality rates that are representative for a large sample of mice. We found an overall incidence rate of 60% and a mortality rate of 20%. While both incidence rates and mortality rates vary significantly in between labs (as evidenced by the high standard deviations when calculating dissecting AAA incidence on a study level, Table 1), we think that these numbers, along with the confounding factor-dependent incidence rates provided in Table 2, can provide an important reference for future researchers planning to use this mouse model. The incidence rates of Grade I-Grade IV morphology should be interpreted with caution as these categories have not been defined very strictly. When first introducing the subdivision into Grade I – Grade IV, Daugherty *et al.* already stated to have "arbitrarily defined aneurysms based largely on the visual characteristics" [20]. Especially the Grade IV category is to be interpreted with caution, as some authors automatically include mice succumbing to aneurysm rupture as Grade IV, while others do not. Nevertheless, we believe that our data are the first to gather representative statistics on the

distribution of these morphological features of murine dissecting AAA, and can as such be important for the interpretation of future results.

When it comes to confounding factors influencing dissecting AAA, we found that incidence and mortality in angiotensin II-infused mice are much higher in male than in female mice. This result confirms earlier findings [37], and is also in correspondence with what is observed in human aneurysms [1, 38]. Our analysis also confirms earlier reports that C57Bl\6 mice have lower dissecting AAA incidence and mortality than ApoE[-/-] mice, while C57Bl\6 that are injected with anti-TGF-β have a significantly higher dissecting AAA incidence and mortality [39]. The reason behind these observations is the subject of ongoing discussion [40, 41].

We further demonstrate that, while a lower dose reduces incidence and mortality of dissecting AAAs, increasing the dose of angiotensin II plays a less important role in incidence than what is commonly assumed [42]. Even when the threshold was fixed at very high levels (> 2000 ng/kg/min), a high dose of angiotensin II did not significantly increase dissecting AAA incidence nor mortality. This is important information that may be used when planning future experiments.

Age turned out not to be related to the incidence of dissecting AAA to the same extent as sex, background and dose. While incidence was significantly different for young and old mice, mortality was not. Moreover, the odds ratio for dissecting AAA incidence was less outspoken, and the confidence interval closer to 1, than for other confounding factors. The thresholds to define young (< 9 weeks) and old (> 24 weeks) mice were based on biological evidence[43], allowing for a reasonable distribution between categories (Figure 2). However, if the age limits were put at 10 weeks and 20 weeks, neither young nor old mice showed a significant difference in dissecting AAA incidence (or mortality). This indicates that the dependency of dissecting

AAA incidence on age is less robust than its dependency on, for instance, sex or angiotensin II dose.

Surprisingly, only moderate and non-significant effects were noted for mice on a high-fat diet. The reason might be that we treated high-fat diet as a binary variable and did not include the amount of cholesterol into the analysis. Most likely a significant difference would have been obtained if a dose-dependent analysis (such as for age and angiotensin II dose) had been performed. The results of the analysis regarding high-fat diet should therefore be interpreted with caution.

The enhancement of reductions and the reduction of enhancements: a short analysis

The most surprising result of our analysis was the strong dependence on study design that was found in both incidence and mortality of dissecting AAA in non-treated, angiotensin II-infused mice that were used as control animals in treatment studies. In part, this can be explained by an intentional selection bias, as researchers select animals with a low incidence (e.g. low dose, wild type mice) when the aim of their study is to enhance dissecting AAA, and focus on animals at higher risk (e.g. male mice or high doses of angiotensin II) when the aim of the study is to reduce dissecting AAA (Figure 2). Nevertheless, a stratified analysis demonstrated that the design-dependency of aneurysm incidence was independent of genetic background, sex or dose of angiotensin II. On the other hand, the design-dependency of mortality was found to normalize after stratification (Figure 3). But while mortality is a binary variable that, in principle, cannot be interpreted wrongly, dissecting AAA incidence is a measurement that is prone to interpretation.

We therefore hypothesized that a possible explanation for the dependency of dissecting AAA incidence on study design could be found in a detection bias with respect to the definition (and the interpretation of that definition) that was used to quantify dissecting AAA presence. Indeed,

17

a thorough analysis revealed the existence of no less than 30 possibilities to quantify the dilatation related to aneurismal widening of the suprarenal segment of the aorta. On top of this, 16% of the studies did not quantify anything at all, as they simply defined dissecting AAA presence after visual inspection of the post mortem tissues. Figure 4 confirms that, while the incidence of dissecting AAA is significantly different for several measurement related factors, mortality is not. The additional stratification in Figure 5 suggests that the dependency of dissecting AAA incidence on study design was to a great extent determined by the measurement method that was used. These data are further supported by the results of the Fisher exact test (Table 3), which demonstrate that both the reference to which diameters are compared and the threshold in percentage increase used to define dissecting AAA are significantly correlated with study design. The results of these stratified analyses should be interpreted with caution due to the relatively low number of animals remaining after stratification for measurement-related factors. Nevertheless, this hypothesis is worrying, as it suggests that dissecting AAA incidence may depend on the method that was used to quantify it.

Another possible explanation for the dependency of incidence and mortality on study design is publication bias. If the control group in an experiment designed to reduce dissecting AAA has a high incidence of dissecting AAA, then the treatment group is more likely to significantly reduce that incidence. Conversely, when the control group reveals an average incidence, then it is more likely that the incidence in the treatment group will not be significantly lower. Vice versa, if the control group in an experiment designed to enhance dissecting AAA has a low incidence, the treatment is more likely to be effective. Since negative results are more difficult to publish, many of these studies will never be submitted or accepted in academic journals, and will thus not be accounted for in the meta-analysis. Such publication bias might be the reason why the incidence and mortality rates of dissecting AAA in our meta-analysis were significantly different for treatment and descriptive studies. While hard to verify, this hypothesis is equally

worrying as it suggests that many of the statistically significant results in published treatment studies were, in fact, chance findings.

A third hypothesis is that the dissecting AAA incidence and mortality rates suffered from a systematic bias between treatment and control groups that was related to the design of the studies but was not included in the stratification. As discussed in depth by Krauth *et al.* [44], additional sources of bias in animal studies (other than the ones already mentioned above) include attrition bias (i.e. not accounting for all animals included in the study), non-randomized treatment allocation, non-blinded analysis, the use of animals with comorbidity, non-compliance with animal welfare requirements, the use of inappropriate statistical models, an incorrect sample size calculation or financial conflicts of interest. We have not stratified for these potential sources of errors as (i) unlike for selection bias, detection bias or publication bias, there was no clear hypothesis to link these factors to study design, and (ii) each additional stratification step would reduce the number of remaining animals, and hence further increase the size of the confidence intervals. Nevertheless, it is important to take these potential sources of error into account when planning future experiments.

Possible solutions and suggestions for future research

When interpreting the results of pre-clinical studies investigating the effect of pharmacological or genetic treatment on dissecting AAA incidence, one often implicitly assumes that the latter is an unequivocal, straightforward observation that is not susceptible to interpretation. Reality, however, is different. In a clinical setting, a luminal increase of 150 % in aortic diameter is typically used as the cut-off criterion to define AAA incidence. This is an artificial limit that is based on the clinical reality that smaller dilatations are at a lower risk of rupture and should therefore not be treated since the surgery risk would outweigh the rupture risk [45]. Despite the fact that treatment studies in a pre-clinical research setting function within an entirely different paradigm, the definition of dissecting AAA has not been adjusted accordingly. Already in 2007,

Jiang *et al.* argued that the "pseudoaneurysms" induced by angiotensin II are a binary event, and that measures of the luminal or external aortic diameter might not accurately reflect the development of murine dissecting AAA, especially when studying the effect of drugs on dissecting AAA incidence and severity [46]. Here, we suggest two different strategies for both the quantification and interpretation of dissecting AAA incidence in studies using the angiotensin II-infused ApoE$^{-/-}$ mouse model:

(i)   When analyzing data, future researchers should refrain from reporting either external or luminal diameters as if they were a single and non-biased expression of dissecting AAA incidence. Instead, the effect of treatment studies should be quantified on a micro-structural level, quantifying e.g. the elastin and collagen content, the size of the intramural thrombus, or the number of branches affected.

(ii)   When interpreting data, future researchers should keep the possibility of publication bias into account, and compare incidence and mortality rates of the angiotensin II-infused control group to the reference values provided in this meta-analysis.

Limitations

The data extraction for the meta-analysis has been performed by a single operator (B.T.). We chose to maximize the number of animals by including those manuscripts in which aneurysm incidence or mortality rates were not mentioned in written text, but only in the form of bar plots, or relative to the group size. Therefore some of the values on dissecting AAA incidence and mortality mentioned in table S1 may be prone to interpretation errors, while others may have gone undetected while they were mentioned within the manuscript. Given the sample size that was achieved, such errors are expected to be randomly distributed and should not affect the conclusions of our work. We would like to point out that the number of studies reporting incidence values was in general greater than those reporting mortality values. In order not to

lose a significant amount of data prior to stratification, the latter was performed for the largest possible groups in both cases. A more stringent interpretation, only taking into account those studies that reported both incidence and mortality, resulted in larger confidence intervals, but did not change the interpretation of the results for dissecting AAA incidence. We did not report these results as the odds ratios for mortality became more difficult to interpret, due to the low number of studies remaining after stratification. On a similar note, the wide confidence intervals after stratification for the measurement methods indicate that the number of mice remaining in this cohort after many subsequent stratification steps was quite low (Figure 5, Table 4). While the exact numerical values of this part of the stratification should therefore be interpreted with caution, they do not alter the main conclusions, which are supported by the Fisher exact test that was carried out on a study level (Table 3).

Conclusions

We have performed a meta-analysis to provide overall as well as stratified, confounding factor-dependent incidence, severity and mortality rates for dissecting AAA in angiotensin II-infused mice. We strongly believe that these numbers have the potential to serve as reference data and will allow for a more accurate study design when planning experiments using the angiotensin II-infused mouse model. The reported incidence and mortality rates were higher for angiotensin II-infused mice that served as control animals in studies designed to reduce dissecting AAA, and lower for mice that served as control animals in studies designed to enhance dissecting AAA. Stratification showed that this dependency could be explained by an intentional selection bias of environmental confounding factors (sex, dose, genetic background, age, diet) for dissecting AAA mortality, but not for dissecting AAA incidence. We subsequently hypothesized that the dependency of dissecting AAA incidence on study design may be related to the variability in methods used to quantify dissecting AAA, and that literature may be

distorted by a publication bias. We conclude that, before any further treatment studies are performed, the translational aspects of this mouse model should be further documented and the relevance of diameter increase as a single quantification of dissecting AAA incidence should be questioned.

Funding

Acknowledgments

Conflict of Interest

None declared.

References

1. Sakalihasan N, Limet R, Defawe OD. Abdominal aortic aneurysm. *Lancet* 2005;**365**:1577-1589.
2. Daugherty A, Cassis LA. Mouse models of abdominal aortic aneurysms. *Arterioscler Thromb Vasc Biol* 2004;**24**:429-434.
3. Daugherty A, Manning MW, Cassis LA. Angiotensin II promotes atherosclerotic lesions and aneurysms in apolipoprotein E-deficient mice. *J Clin Invest* 2000;**105**:1605-1612.
4. Rateri DL, Howatt DA, Moorleghen JJ, Charnigo R, Cassis LA, Daugherty A. Prolonged Infusion of Angiotensin II in apoE–/– Mice Promotes Macrophage Recruitment with Continued Expansion of Abdominal Aortic Aneurysm. *Am J Pathol* 2011;**179**:1542-1548.
5. Saraff K, Babamusta F, Cassis LA, Daugherty A. Aortic dissection precedes formation of aneurysms and atherosclerosis in angiotensin II-infused, apolipoprotein E-deficient mice. *Arterioscler Thromb Vasc Biol* 2003;**23**:1621-1626.
6. Ahluwalia N, Lin AY, Tager AM, Pruitt IE, Anderson TJT, Kristo F, Shen D, Cruz AR, Aikawa M, Luster AD, Gerszten RE. Inhibited Aortic Aneurysm Formation in BLT1-Deficient Mice. *J Immunol* 2007;**179**:691-697.
7. Bhamidipati CM, Whatling CA, Mehta GS, Meher AK, Hajzus VA, Su G, Salmon M, Upchurch GR, Owens GK, Ailawadi G. 5-Lipoxygenase Pathway in Experimental Abdominal Aortic Aneurysms. *Arterioscler Thromb Vasc Biol* 2014;**34**:2669-2678.
8. Deguchi J-o, Huang H, Libby P, Aikawa E, Whittaker P, Sylvan J, Lee RT, Aikawa M. Genetically engineered resistance for MMP collagenases promotes abdominal aortic aneurysm formation in mice infused with angiotensin II. *Lab Invest* 2009;**89**:315-326.
9. Gitlin JM, Trivedi DB, Langenbach R, Loftin CD. Genetic deficiency of cyclooxygenase-2 attenuates abdominal aortic aneurysm formation in mice. *Cardiovasc Res* 2007;**73**:227-236.
10. Yokoyama U, Ishiwata R, Jin M-H, Kato Y, Suzuki O, Jin H, Ichikawa Y, Kumagaya S, Katayama Y, Fujita T, Okumura S, Sato M, Sugimoto Y, Aoki H, Suzuki S, Masuda M, Minamisawa S, Ishikawa Y. Inhibition of EP4 Signaling Attenuates Aortic Aneurysm Formation. *PLoS ONE* 2012;**7**:e36724.
11. Basu R, Fan D, Kandalam V, Lee J, Das SK, Wang X, Baldwin TA, Oudit GY, Kassiri Z. Loss of Timp3 Gene Leads to Abdominal Aortic Aneurysm Formation in Response to Angiotensin II. *J Biol Chem* 2012;**287**:44083-44096.
12. Haldar SM, Lu Y, Jeyaraj D, Kawanami D, Cui Y, Eapen SJ, Hao C, Li Y, Doughman Y-Q, Watanabe M, Shimizu K, Kuivaniemi H, Sadoshima J, Margulies KB, Cappola TP, Jain MK. Klf15 Deficiency Is a Molecular Link Between Heart Failure and Aortic Aneurysm Formation. *Sci Transl Med* 2010;**2**.
13. Kitagawa T, Kosuge H, Chang E, James ML, Yamamoto T, Shen B, Chin FT, Gambhir SS, Dalman RL, McConnell MV. Integrin-Targeted Molecular Imaging of Experimental Abdominal Aortic Aneurysms by 18F-labeled Arg-Gly-Asp Positron-Emission Tomography. *Circ Card Imag* 2013;**6**:950-956.
14. Li F, Downing BD, Smiley LC, Mund JA, DiStasi MR, Bessler WK, Sarchet KN, Hinds DM, Kamendulis LM, Hingtgen CM, Case J, Clapp DW, Conway SJ, Stansfield BK, Ingram DA. Neurofibromin-Deficient Myeloid Cells Are Critical Mediators of Aneurysm Formation In Vivo. *Circulation* 2014;**129**:1213-1224.
15. Merino V, Todiras M, Mori M, Sales VT, Fonseca R, Saul V, Tenner K, Bader M, Pesquero J. Predisposition to atherosclerosis and aortic aneurysms in mice deficient in kinin B1 receptor and apolipoprotein E. *J Mol Med* 2009;**87**:953-963.

16.   Usui F, Shirasuna K, Kimura H, Tatsumi K, Kawashima A, Karasawa T, Yoshimura K, Aoki H, Tsutsui H, Noda T, Sagara J, Taniguchi Si, Takahashi M. Inflammasome Activation by Mitochondrial Oxidative Stress in Macrophages Leads to the Development of Angiotensin II–Induced Aortic Aneurysm. *Arterioscler Thromb Vasc Biol* 2014.

17.   Wang J, Lindholt JS, Sukhova GK, Shi MA, Xia M, Chen H, Xiang M, He A, Wang Y, Xiong N, Libby P, Wang J-A, Shi G-P. IgE actions on CD4(+) T cells, mast cells, and macrophages participate in the pathogenesis of experimental abdominal aortic aneurysms. *EMBO Mol Med* 2014;**6**:952-969.

18.   Cassis LA, Helton MJ, Howatt DA, King VL, Daugherty A. Aldosterone does not mediate angiotensin II-induced atherosclerosis and abdominal aortic aneurysms. *Br J Pharmacol* 2005;**144**:443-448.

19.   Daugherty A, Cassis LA, Lu H. Complex pathologies of angiotensin II-induced abdominal aortic aneurysms. *J Zhejiang Univ (Agric Life Sci)* 2011;**12**:624-628.

20.   Daugherty A, Manning MW, Cassis LA. Antagonism of AT2 receptors augments Angiotensin II-induced abdominal aortic aneurysms and atherosclerosis. *Br J Pharmacol* 2001;**134**:865-870.

21.   Daugherty A, Rateri DL, Howatt DA, Charnigo R, Cassis LA. PD123319 Augments Angiotensin II-Induced Abdominal Aortic Aneurysms through an AT2 Receptor-Independent Mechanism. *PLoS ONE* 2013;**8**:e61849.

22.   Police SB, Putnam K, Thatcher S, Batifoulier-Yiannikouris F, Daugherty A, Cassis LA. Weight loss in obese C57BL/6 mice limits adventitial expansion of established angiotensin II-induced abdominal aortic aneurysms. *Am J Physiol - Heart Circ Physiol* 2010;**298**:H1932-H1938.

23.   Xie X, Lu H, Moorleghen JJ, Howatt DA, Rateri DL, Cassis LA, Daugherty A. Doxycycline Does Not Influence Established Abdominal Aortic Aneurysms in Angiotensin II-Infused Mice. *PLoS One* 2012;**7**:e46411.

24.   Malekzadeh S, Fraga-Silva RA, Trachet B, Montecucco F, Mach F, Stergiopulos N. Role of the renin–angiotensin system on abdominal aortic aneurysms. *Eur J Clin Invest* 2013;**43**:1328-1338.

25.   Saito T, Hasegawa Y, Ishigaki Y, Yamada T, Gao J, Imai J, Uno K, Kaneko K, Ogihara T, Shimosawa T, Asano T, Fujita T, Oka Y, Katagiri H. Importance of endothelial NF-κB signalling in vascular remodelling and aortic aneurysm formation. *Cardiovasc Res* 2013;**97**:106-114.

26.   Barisione C, Charnigo R, Howatt DA, Moorleghen JJ, Rateri DL, Daugherty A. Rapid dilation of the abdominal aorta during infusion of angiotensin II detected by noninvasive high-frequency ultrasonography. *J Vasc Surg* 2006;**44**:372-376.

27.   Cao RY, Amand T, Ford MD, Piomelli U, Funk CD. The Murine Angiotensin II-Induced Abdominal Aortic Aneurysm Model: Rupture Risk and Inflammatory Progression Patterns. *Front Pharmacol* 2010;**1**:9-9.

28.   Gavish L, Beeri R, Gilon D, Rubinstein C, Berlatzky Y, Gavish LY, Bulut A, Harlev M, Reissman P, Gertz SD. Inadequate reinforcement of transmedial disruptions at branch points subtends aortic aneurysm formation in apolipoprotein-E-deficient mice. *Cardiovasc Pathol* 2014;**23**:152-159.

29.   Gavish L, Rubinstein C, Berlatzky Y, Gavish LY, Beeri R, Gilon D, Bulut A, Harlev M, Reissman P, Gertz SD. Low level laser arrests abdominal aortic aneurysm by collagen matrix reinforcement in apolipoprotein E-deficient mice. *Lasers Surg Med* 2012;**44**:664-674.

30.   Prins PA, Hill MF, Airey D, Nwosu S, Perati PR, Tavori H, F. Linton M, Kon V, Fazio S, Sampson UK. Angiotensin-Induced Abdominal Aortic Aneurysms in

Hypercholesterolemic Mice: Role of Serum Cholesterol and Temporal Effects of Exposure. *PLoS One* 2014;**9**:e84517.

31.   Sampson UK, Perati PR, Prins PA, Pham W, Liu Z, Harrell FE, Jr., Linton MF, Gore JC, Kon V, Fazio S. Quantitative Estimates of the Variability of In Vivo Sonographic Measurements of the Mouse Aorta for Studies of Abdominal Aortic Aneurysms and Related Arterial Diseases. *J Ultrasound Med* 2011;**30**:773-784.

32.   Trachet B, Renard M, De Santis G, Staelens S, De Backer J, Antiga L, Loeys B, Segers P. An integrated framework to quantitatively link mouse-specific hemodynamics to aneurysm formation in angiotensin II-infused ApoE -/- mice. *Ann Biomed Eng* 2011;**39**:2430-2444.

33.   Trachet B, Renard M, Van der Donckt C, Deleye S, Bols J, De Meyer GRY, Staelens S, Loeys BL, Segers P. Longitudinal follow-up of ascending versus abdominal aortic aneurysm formation in angiotensin II-infused ApoE–/– mice. *Artery research* 2014;**8**:16-23.

34.   Gavish L, Beeri R, Gilon D, Rubinstein C, Berlatzky Y, Bulut A, Reissman P, Gavish LY, Gertz SD. Arrest of progression of pre-induced abdominal aortic aneurysm in apolipoprotein E-deficient mice by low level laser phototherapy. *Lasers Surg Med* 2014;**46**:781-790.

35.   Trachet B, Fraga-Silva RA, Piersigilli A, Tedgui A, Sordet-Dessimoz J, Astolfo A, Van der Donckt C, Modregger P, Stampanoni M, Segers P, Stergiopulos N. Dissecting abdominal aortic aneurysm in ang II-infused mice: suprarenal branch ruptures and apparent luminal dilatation. *Cardiovasc Res* 2015;**105**:213-222.

36.   Schriefl AJ, Collins MJ, Pierce DM, Holzapfel GA, Niklason LE, Humphrey JD. Remodeling of Intramural Thrombus and Collagen in an Ang-II Infusion ApoE–/– Model of Dissecting Aortic Aneurysms. *Thromb Res* 2012;**130**:e139-e146.

37.   Zhang X, Thatcher S, Wu C, Daugherty A, Cassis LA. Castration of male mice prevents the progression of established angiotensin II-induced abdominal aortic aneurysms. *J Vasc Surg*.

38.   Alexander JJ. The pathobiology of aortic aneurysms. *J Surg Res* 2004;**117**:163-175.

39.   Wang Y, Ait-Oufella H, Herbin O, Bonnin P, Ramkhelawon B, Taleb S, Huang J, Offenstadt G, Combadiere C, Renia L, Johnson JL, Tharaux P-L, Tedgui A, Mallat Z. TGF-beta activity protects against inflammatory aortic aneurysm progression and complications in angiotensin II-infused mice. *J Clin Invest* 2010;**120**:422-432.

40.   Klink A, Heynens J, Herranz B, Lobatto ME, Arias T, Sanders HMHF, Strijkers GJ, Merkx M, Nicolay K, Fuster V, Tedgui A, Mallat Z, Mulder WJM, Fayad ZA. In Vivo Characterization of a New Abdominal Aortic Aneurysm Mouse Model With Conventional and Molecular Magnetic Resonance Imaging. *J Am Coll Cardiol* 2011;**58**:2522-2530.

41.   Ait-Oufella H, Wang Y, Herbin O, Bourcier S, Potteaux S, Joffre J, Loyer X, Ponnuswamy P, Esposito B, Dalloz M, Laurans L, Tedgui A, Mallat Z. Natural Regulatory T Cells Limit Angiotensin II–Induced Aneurysm Formation and Rupture in Mice. *Arterioscler Thromb Vasc Biol* 2013;**33**:2374-2379.

42.   Ayabe N, Babaev VR, Tang Y, Tanizawa T, Fogo AB, Linton MF, Ichikawa I, Fazio S, Kon V. Transiently heightened angiotensin II has distinct effects on atherosclerosis and aneurysm formation in hyperlipidemic mice. *Atherosclerosis* 2006;**184**:312-321.

43.   Finlay BL, Darlington RB. Linked regularities in the development and evolution of mammalian brains. *Science* 1995;**268**:1578-1584.

44.   Krauth D, Woodruff TJ, Bero L. Instruments for Assessing Risk of Bias and Other Methodological Criteria of Published Animal Studies: A Systematic Review. *Environ Health Perspect* 2013;**121**:985-992.

45.   Mortality results for randomised controlled trial of early elective surgery or ultrasonographic surveillance for small abdominal aortic aneurysms. The UK Small Aneurysm Trial Participants. *Lancet* 1998;**352**:1649-1655.

46.   Jiang F, Jones GT, Dusting GJ. Failure of antioxidants to protect against angiotensin II-induced aortic rupture in aged apolipoprotein(E)-deficient mice. *Br J Pharmacol* 2007;**152**:880-890.

Figures
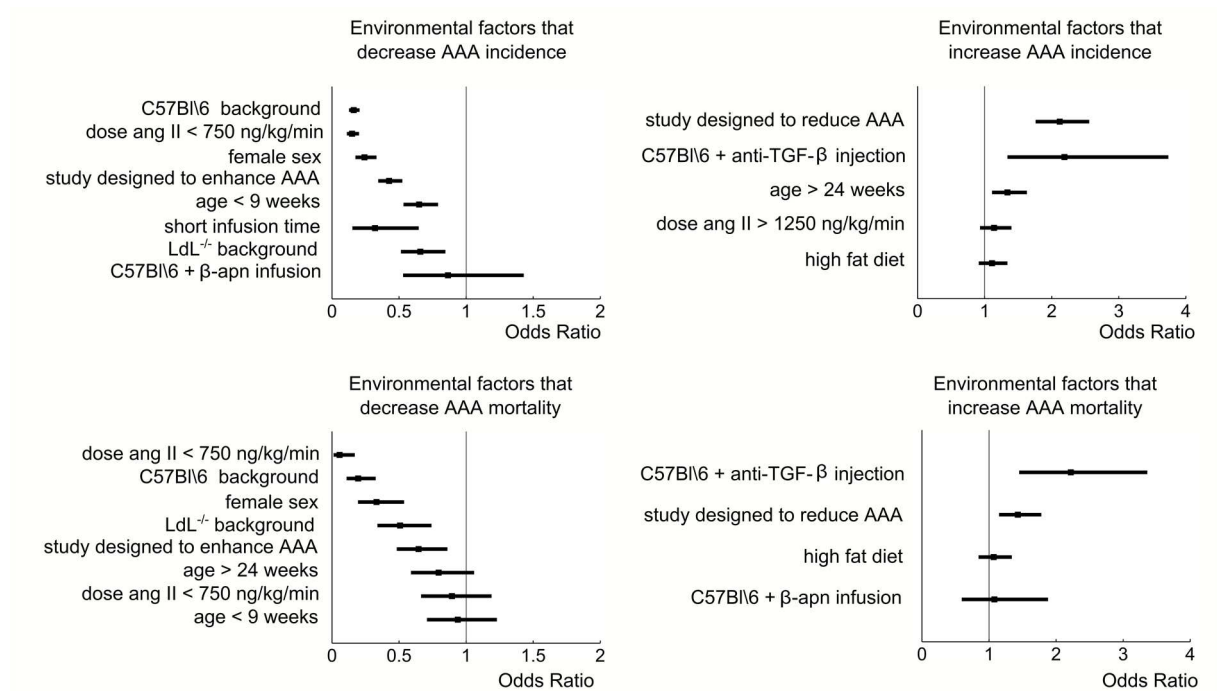
Figure 1. Odds ratios and 95% confidence intervals for environmental confounding factors. Every factor is compared to its reference value (male sex, ApoE $^{-/-}$ background, age 9-24 weeks, normal diet, dose 750 - 1250 ng/kg/min , long infusion time, study designed to describe dissecting AAA). Odds ratios are considered significant if the confidence interval does not contain the value 1. Factors are shown in order of increasing p-value.
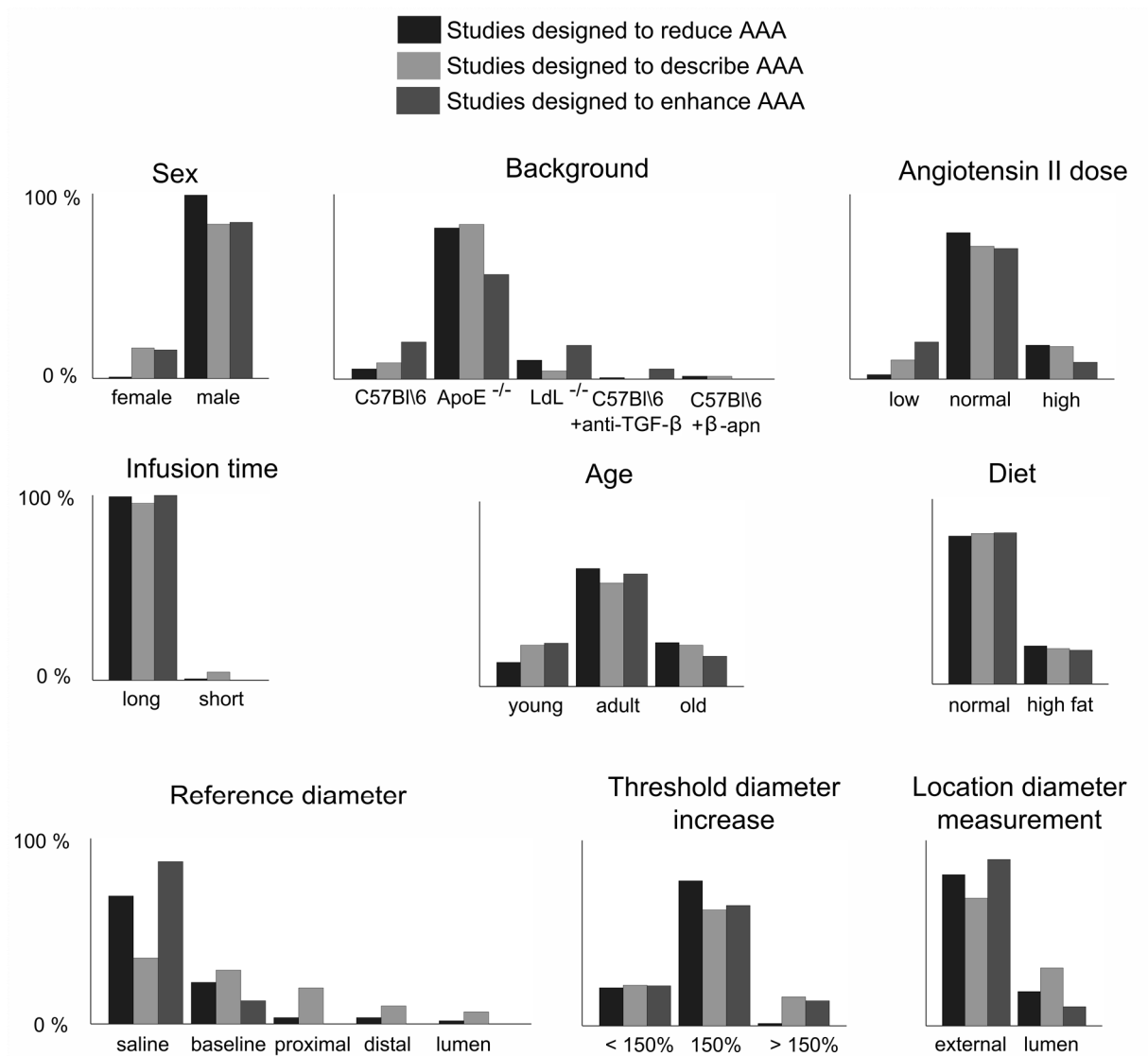
Figure 2. Histograms indicating the prevalence of each category for both environmental confounding factors (top 6 panels) and measurement-related factors (bottom 3 panels) on a study level. Histograms are subdivided for 3 different study designs and normalized by the number of studies, which can be found in Table 3.

Figure 3. Step-wise stratification to investigate whether environmental confounding factors can explain the dependency of dissecting AAA incidence and mortality on study design. Factors were included in decreasing order of significance of their correlation with study design, as determined from Fisher's exact test on a study level (see Table 3). Top panels show that dissecting AAA incidence remains significantly different after stratification, while bottom panels show that dissecting AAA mortality becomes independent of study design after stratification.
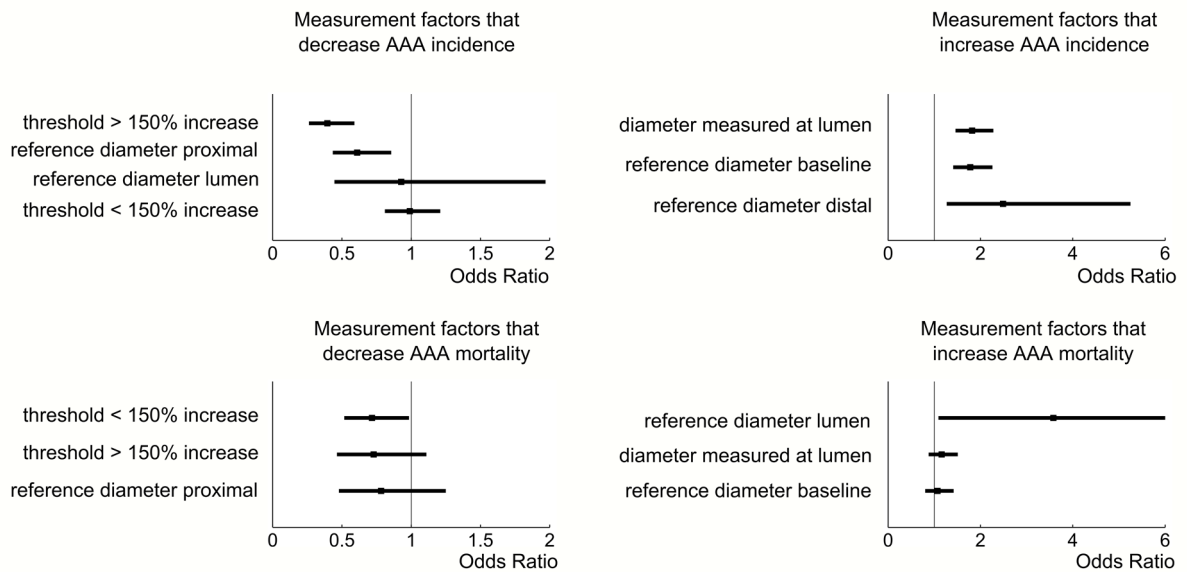
Figure 4. Odds ratios and 95% confidence intervals for measurement related factors. Every factor is compared to its reference value (a 150% increase in dissecting AAA external diameter, with a reference diameter obtained from the same supra-renal segment in saline-infused mice). Odds ratios are considered significant if the confidence interval does not contain the value 1.
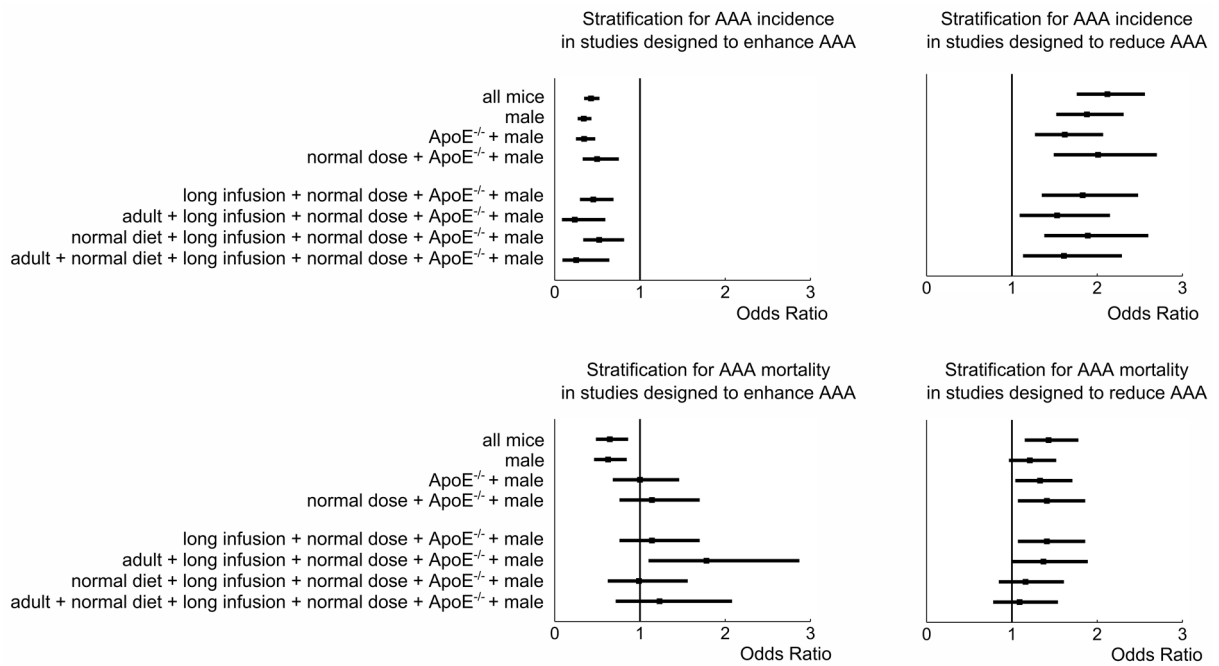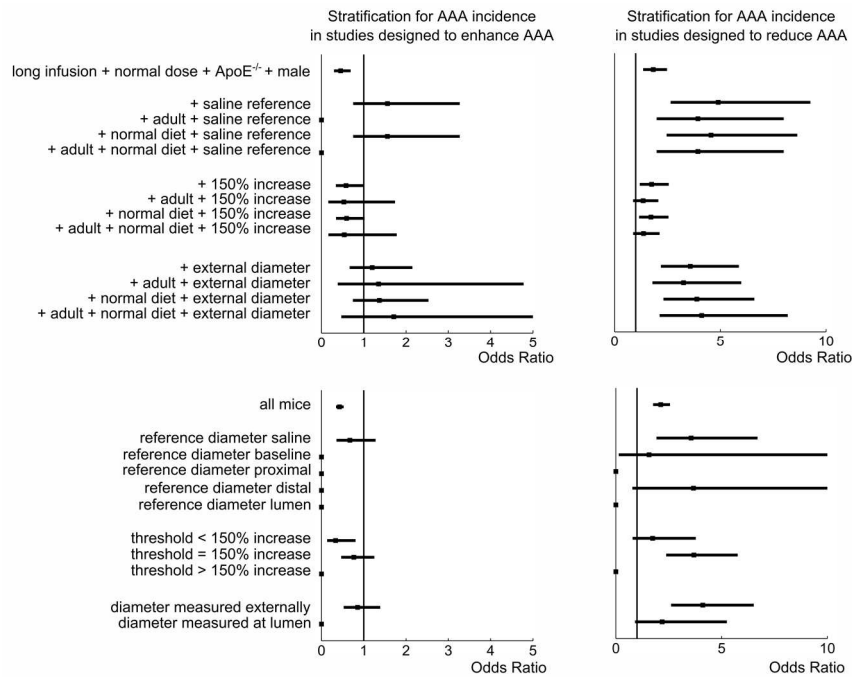
Figure 5. Step-wise stratification to investigate whether measurement-related confounding factors can explain the dependency of dissecting AAA incidence and mortality on study design. **Top.** The starting point are reference mice with additional stratification for factors significantly correlated with study design (sex, background, dose) and infusion time. The reference values of each measurement related factor (saline reference, 150% increase, external measurement) were included in decreasing order of significance of their correlation with study design, as determined from Fisher's exact test on a study level (see Table 3). The effect of including age and diet was investigated for each factor as well. **Bottom.** The starting point are all mice, without any prior stratification. The influence of stratification is subsequently shown for all values of each measurement method.

Tables

	N studies	N mice	N mice Total	Incid. (%) mice	Incid. (%) studies	Std. (%) studies
Incidence	187	1977	3320	60	59	31
Mortality	129	573	2865	20	18	19
Grade I	33	93	429	22	28	28
Grade II	33	110	429	26	26	24
Grade III	33	125	429	29	27	22
Grade IV	33	101	429	24	18	20

Table 1. Dissecting AAA incidence, severity and mortality for all studies included in the meta-analysis. Incidence values were obtained both on an individual mouse level (Incid. mice), taking into account the amount of mice in each study, or on a study level (Incid. studies), averaging incidence values of all studies, irrespective of the amount of mice in each study. Std: standard deviation.

Environmental factors	OR Incid.	95% C.I. Incid.	% Incid. ref	% Incid. factor	NsI ref	NsI factor	OR Mort.	95% C.I. Mort.	% Mort. ref	% Mort. factor	NsM ref	NsM factor
C57Bl\6 background	0.16	[0.13 - 0.21]	66	24	137	24	0.20	[0.11 - 0.33]	22	5	90	16
Ang II dose < 750 ng/kg/min	0.15	[0.11 - 0.20]	63	20	140	17	0.06	[0.01 - 0.17]	22	2	100	12
Female sex	0.24	[0.17 - 0.33]	62	29	163	15	0.33	[0.19 - 0.54]	21	8	107	15
Studies designed to enhance AAA	0.43	[0.34 - 0.52]	56	35	44	43	0.65	[0.48 - 0.86]	19	13	32	31
Studies designed to reduce AAA	2.12	[1.76 - 2.56]	56	73	44	100	1.43	[1.15 - 1.78]	19	25	32	66
Age < 9 weeks	0.65	[0.53 - 0.79]	61	50	103	27	0.94	[0.71 - 1.23]	21	20	72	17
Ang II infusion Time < 7 days	0.32	[0.15 - 0.65]	60	33	185	4	N/A	[0 - Inf]	20	N/A	130	0
C57Bl\6 + anti-TGF-β injection	2.19	[1.34 - 3.74]	66	81	137	3	2.22	[1.45 - 3.36]	22	39	90	3
LdL-/- background	0.66	[0.51 - 0.85]	66	56	137	21	0.51	[0.34 - 0.74]	22	13	90	18
Age > 24 weeks	1.34	[1.11 - 1.63]	61	68	103	43	0.80	[0.59 - 1.06]	21	18	72	26
Dose Ang II > 1250 ng/kg/min	1.14	[0.93 - 1.40]	63	66	140	31	0.89	[0.66 - 1.19]	22	20	100	17
High fat diet	1.11	[0.91 - 1.34]	59	62	154	35	1.07	[0.84 - 1.34]	20	21	104	26
C57Bl\6 + β-apn	0.87	[0.53 - 1.43]	66	62	137	3	1.08	[0.59 - 1.88]	22	23	90	3
Measurement factors												
Diameter measured at lumen	1.82	[1.46 - 2.28]	59	72	92	20	1.16	[0.88 - 1.51]	19	21	57	14
Reference diameter baseline	1.78	[1.41 - 2.26]	59	72	66	20	1.07	[0.80 - 1.42]	20	21	44	14
Threshold > 150% increase	0.40	[0.26 - 0.59]	62	39	95	10	0.73	[0.46 - 1.11]	21	16	58	9
Reference diameter proximal	0.61	[0.43 - 0.87]	59	47	66	7	0.78	[0.48 - 1.25]	20	16	44	6
Reference diameter distal	2.49	[1.27 - 5.25]	59	78	66	5	N/A	[0 - Inf]	20	N/A	44	0
Reference diameter lumen	0.93	[0.45 - 1.97]	59	57	66	2	3.58	[1.09 - 11.50]	20	47	44	1
Threshold < 150% increase	0.99	[0.81 - 1.21]	62	62	95	27	0.72	[0.52 - 0.98]	21	16	58	12

Table 2. Odds ratios and 95% confidence intervals for incidence rates (left) and mortality rates (right) of dissecting AAA in association with environmental confounding factors (top) and measurement-related factors (bottom). Every factor is compared to its reference value (male sex, ApoE $^{-/-}$ background, age 9-24 weeks, normal diet, dose 750-1250 ng/kg/min, infusion time > 6 days, study designed to describe dissecting AAA, diameter measured externally, dissecting AAA defined as a diameter increase of 150%, in comparison with saline-infused reference animals). Odds ratios are considered significant if the 95% confidence interval does not contain the value 1. Factors are listed in order of increasing p-value for the odds ratio of dissecting AAA incidence.

Incid.: Incidence, Mort.: Mortality, OR: Odds Ratio, CI: Confidence Interval, ref: reference, NsI: number of studies reporting incidence, NsM: number of studies reporting mortality.

	NsR	NsD	NsE	p-value Fs	p-value Fm incid.	p-value Fm mort.
Environmental factors						
Sex	120	61	52	2.2E-05	8.3E-54	7.7E-07
Background	120	62	44	8.2E-04	N\A	1.9E-26
Ang II dose	126	68	55	1.7E-03	2.0E-72	6.5E-23
Infusion time	127	68	55	1.2E-01	1.7E-06	1.0E+00
Age	123	63	43	3.6E-01	2.2E-10	2.0E-05
Diet	127	68	55	9.6E-01	1.3E-05	3.2E-02
Measurement factors						
Reference diameter	31	58	24	2.7E-03	N\A	3.7E-29
Threshold increase	78	32	37	2.4E-02	2.7E-36	1.2E-35
Diameter location	65	32	29	1.3E-01	7.3E-22	4.2E-14

Table 3. Results of Fisher's exact test indicate to what extent each confounding factor is related to study design. Fisher's exact test was computed both on a study level (Fs) and for dissecting AAA incidence and mortality on an individual mouse level (Fm). Corresponding histograms on a study level are shown in Figure 2, and the corresponding p-values determined the order of stratification in Figures 3 and 5.

NsR: number of studies designed to reduce dissecting AAA, which report the indicated factor. NsD: number of studies designed to describe dissecting AAA, which report the indicated factor. NsE: number of studies designed to enhance dissecting AAA, which report the indicated factor. Fs: Fisher's exact test on a study level. Fm: Fisher's exact test on an individual mouse level. Incid.: incidence. Mort.: Mortality. N\A: not available. Since these factors take 5 different values, 3 of which are very infrequent, the Fisher's exact test could not be computed on an individual mouse level.

AAA incidence	OR reduce	95% C.I. reduce	OR enhance	95% C.I. enhance	NsD	NsR	NsE
male	1.88	[1.52 - 2.31]	0.34	[0.27 - 0.43]	32	94	35
ApoE -/- + male	1.62	[1.27 - 2.07]	0.35	[0.25 - 0.48]	27	75	14
normal dose + ApoE -/- + male	2.01	[1.49 - 2.70]	0.50	[0.33 - 0.75]	20	57	9
long infusion + normal dose + ApoE -/- + male	1.83	[1.35 - 2.48]	0.45	[0.30 - 0.69]	17	57	9
adult + long infusion + normal dose + ApoE -/- + male	1.53	[1.09 - 2.15]	0.24	[0.08 - 0.59]	14	42	2
normal diet + long infusion + normal dose + ApoE -/- + male	1.89	[1.38 - 2.60]	0.52	[0.33 - 0.81]	14	52	8
adult + normal diet + long infusion + normal dose + ApoE -/- + male	1.61	[1.13 - 2.29]	0.25	[0.09 - 0.64]	11	38	2
Stratified + saline reference	4.90	[2.65 - 9.26]	1.56	[0.75 - 3.27]	6	20	3
Stratified + 150% increase	1.75	[1.19 - 2.56]	0.58	[0.34 - 0.99]	8	38	4
Stratified + saline reference + 150% increase	3.92	[1.72 - 9.16]	1.24	[0.50 - 3.17]	2	20	3
Stratified + saline reference + adult	3.94	[1.99 - 8.00]	0.00	[0 - Inf]	5	13	0
Stratified + 150% increase + adult	1.35	[0.87 - 2.07]	0.53	[0.16 - 1.74]	6	29	1
Stratified + saline reference + 150% increase + adult	3.05	[1.13 - 8.60]	0.00	[0 - Inf]	1	13	0
Stratified + saline reference + normal diet	4.56	[2.46 - 8.64]	1.56	[0.75 - 3.27]	6	19	3
Stratified + 150% increase + normal diet	1.72	[1.16 - 2.55]	0.59	[0.35 - 1.02]	6	36	4
Stratified + saline reference + 150% increase + normal diet	3.65	[1.59 - 8.54]	1.24	[0.50 - 3.17]	2	19	3
Stratified + saline reference + adult + normal diet	3.94	[1.99 - 8.00]	0.00	[0 - Inf]	5	13	0
Stratified + 150% increase + adult + normal diet	1.37	[0.88 - 2.13]	0.54	[0.16 - 1.78]	4	28	1
Stratified + saline reference + 150% increase + adult + normal diet	3.05	[1.13 - 8.60]	0.00	[0 - Inf]	1	13	0
reference diameter saline	3.56	[1.92, 6.70]	0.67	[0.36, 1.28]	4	13	7
reference diameter baseline	1.57	[0.14, 11.60]	0.00	[0, Inf]	1	6	0
reference diameter proximal	0.00	[0, Inf]	0.00	[0, Inf]	2	0	0
reference diameter distal	3.67	[0.78, 17.80]	0.00	[0, Inf]	3	2	0
reference diameter lumen	0.00	[0, Inf]	0.00	[0, Inf]	1	0	0
threshold < 150% increase	1.74	[0.798, 3.78]	0.33	[0.14, 0.81]	2	10	4
threshold = 150% increase	3.69	[2.38, 5.76]	0.77	[0.47, 1.25]	10	20	8
threshold > 150% increase	0.00	[0, Inf]	0.00	[0, Inf]	0	1	1
diameter measured externally	4.11	[2.61, 6.52]	0.86	[0.53, 1.39]	9	19	9
diameter measured at lumen	2.19	[0.90, 5.25]	0.00	[0, Inf]	2	5	0
AAA mortality							
male	1.21	[0.97 - 1.52]	0.626	[0.46 - 0.84]	20	62	24
ApoE -/- + male	1.33	[1.04 - 1.71]	1	[0.68 - 1.46]	18	47	8
normal dose + ApoE -/- + male	1.41	[1.07 - 1.86]	1.14	[0.76 - 1.70]	14	36	7
long infusion + normal dose + ApoE -/- + male	1.41	[1.07 - 1.86]	1.14	[0.76 - 1.70]	14	36	7
adult + long infusion + normal dose + ApoE -/- + male	1.37	[1.00 - 1.89]	1.78	[1.10 - 2.87]	11	27	3
normal diet + long infusion + normal dose + ApoE -/- + male	1.16	[0.85 - 1.61]	0.988	[0.62 - 1.56]	12	34	5
adult + normal diet + long infusion + normal dose + ApoE -/- + male	1.09	[0.78 - 1.54]	1.23	[0.71 - 2.08]	10	27	2

Table 4. Stratification for environmental confounding factors and measurement-related factors. Factors were included in decreasing order of significance of their correlation with study design, as determined from Fisher's exact test (see Table 3). The 'stratified' case refers to long infusion + normal dose + ApoE$^{-/-}$ + male mice. Odds ratios are considered significant if the confidence interval does not contain the value 1.

NsD: number of studies designed to describe dissecting AAA, NsR: number of studies designed to reduce dissecting AAA, NsE: number of studies designed to enhance dissecting AAA.