Review of assignment O1-2016

Eva Ekblad, Professor
Unit Neurogastroenterology,
Department of Experimental Medical Science,
Lund University
BMC B11, SE-22184 Lund, Sweden, email: eva.ekblad@med.lu.se

This is a translation of "Rapport, granskning i ärende O1-2016". In case of discrepancies the Swedish version has preferential interpretation.

Summary

Background:
The scientific article "Experimental orthotopic transplantation of a tissue-engineered oesophagus in rats", Sjöqvist et al, Nature Communications, 5:3562, 2014, was in the same year reported to the Vice-Chancellor of the Karolinska institute (KI) with the suspicion of scientific misconduct. The, by KI appointed, external reviewer concluded that the principal investigator and author of this article was guilty of scientific misconduct on one item. The principal author was later cleared of suspicion by the then present Vice-Chancellor. A new investigation on the suspicions of scientific misconduct was launched in 2016 by the deputy Vice-Chancellor at KI. The Central Ethical Review Board (CEPN) was assigned and they appointed me as an external reviewer of this article and associated documents. Additional documents were also requested.

Assessment:

Based on "The European Code of Conduct for Research Integrity" I find the principal investigator and author guilty of scientific misconduct due to:

- Refusal/inability to disclose all requested raw data for the results presented in the article.
- Refusal/inability to disclose animal log records (försöksdjurjournaler).
- Misleadingly presented, interpreted and described the results.
- Severely deviated from the animals ethics permit.
- Deceiving the licensing authority; the regional animal ethics committee.

Lund 2016-07-09

Content

The assignment

The Expert Group for Misconduct in Research at the Central Ethical Review Board (CEPN) assigned me in March 2016 as external reviewer with the task of investigating suspected scientific misconduct concerning the article ”Experimental orthotopic transplantation of a tissue-engineered oesophagus in rats”, Sjöqvist et al, Nature Communications, 5:3562, 2014. If the results presented were misleading together with the authors' responsibility for the content of the article before and after publication were also to be enlighted.

The background is KI´s and deputy Vice Chancellor K Dahlman-Wrights request to The Expert Group for Misconduct in Research to make a statement on if what is revealed in the matter is scientific misconduct.

The case: A notification from O Simonsson (MD and PhD student), M Corbascio (MD and ass prof) and K-H Grinnemo (MD and researcher) concerning suspected scientific misconduct within the article ”Experimental orthotopic transplantation of a tissue-engineered oesophagus in rats”, by Sjöqvist et al.

Documents

Attached with the assignment from KI:
1. The article, Sjöqvist et al Nature Communications 5:3562, 2014
2. The notification of suspected scientific misconduct 2014-06-24
3. Comments from Paolo Macchiarini (PM) 2014-08-03
4. Sjöqvist lab books (11 files)
5. External review by Professor Bengt Gerdin (BG) 2015-05-13
6. Comments from PM and co-authors, June 2015
7. KI decision, 2015-08-28
8. Prof Gerdins close-up comments, Dec 2015
9. Correspondence PM och Nature Communications

2016-03-08 additional documents requested:
Animal log records (försöksdjursjournaler; according to SJVFS 2015:38, L150: chapter 8) concerning all animals used within the study (n=94, ”methods, animal care and organ harvesting”, the article, page 12).
Copies of the ethics permit(s) granted to conduct the study, including all addendums.

2016-03-17 from KIs administration, Mats Gustavsson and on 2016-03-30 from PM arrival of the same dropbox link containing:
Correspondence between PM and Nature Communications.
Copy of animal ethics permit, S12-13 (not complete since appendix is missing), and one addendum, S36-13, granted afterwards.
Protocols on transplanted animals, body weights and health status – but no animal log records.

2016-04-13 Another dropbox link from PM containing documentation and the original files from the CT scan figure (fig 8a) arrived. In this dropbox also ”appendix to the animal ethics permit diary no S12-13” was included.

2016-04-21 Another dropbox link from KI containing a report from the County Administrative Board, 2016-03-09: control according to Sw. Animal Welfare Act and copies

of ethics permits S175-09, S160-11, S131-11, S110-11, S36-13, S23-14 and S12-13, check lists from the Approved Manager of the animal facility concerning anaestesia for 10 rats (operated on June-November 2013) and photographs of lab protocols (not the animal log records) for 11 rats (operated March-May, year: unknown).

2016-04-22 at a meeting with CEPN the decision to send yet another request for documents together with a couple of questions for clarification was taken. On 2016-04-27 was ”Questions asked and data requested concerning O1-2016” sent from CEPN to PM by way of A Gustafsson, KI with a request to answer before 2016-05-04.

2016-05-16 PM replied to the above-mentioned request.

The article

Experimental orthotopic transplantation of a tissue-engineered oesophagus in rats by S Sjöqvist, P Jungebluth, ML Lim, JC Haag, Y Gustafsson, G Lemon, S Baiguera, MA Burguillos, C del Gaudio, A B Rodriguez, A Sotnichenko, K Kublickiene, H Ullman, H Kielstein, P Damberg, A Bianco, R Heuchel, Y Zhao, D Ribatti, C Ibarra, B Joseph, DA Taylor and P Macchiarini. Published 15th April, 2014 in Nature Communications (5:3562, DOI 10.1038/ncomms4562). Corresponding author: Paolo Macchiarini (PM).

Summary:
The aim of the study is to develop a biological scaffold from oesophagus that is immunologically inactive and that can be repopulated by stem cells from bone marrow with the goal to regenerate a, functionally and structurally, well-functioning transplant for oesophagus. All experimentation was performed on rat and began with *in vitro* investigations on the capacity of the candidate transplant to regenerate the different layers normally present in oesophagus, followed by tests on mechanic and bioelectric properties. The numbers of animals/transplants examined in these various studies are unclear. Immunological properties of the candidate transplant were examined *in vivo* by placing it subcutaneously for five days followed by examinations concerning rejection and presence of immune cells (macrophages). For the latter examination 3 different groups of animals were used with 3 animals in each group.

The last part of the study involves replacing a 15 mm long segment of proximal (upper portion) oesophagus with a 15 mm long transplant prepared with stem cells. Rats with this operation were compared with sham operated rats on which the same portion of oesophagus was removed but resutured in place. The animals were sacrificed 14 days post surgery. They were fed a liquid diet on day 1-7 and thereafter food with soft texture. The number of animals operated on in this way is stated to be 10; how they were distributed between the two groups is not indicated. All animals are reported to have survived the surgery and the 14 postoperative days. They are all reported to have recovered well, to be without pain and at good health. They had however, a significant weight loss; approximately 40% loss of body weight after 14 days, in both groups. In a Supplement a video displays transplanted rats in a relatively good health status; at what postoperative day the video is recorded is not indicated.

The authors conclude that the developed transplant showed structural and functional characteristics comparable to native oesophagus, that it is immunologically inactive and that it contains newly formed blood vessels and nerve fibres. They conclude that the biological scaffold reseeded with stromal cells is ideal and that it fulfills all necessary demands

for use as a graft since it *in vivo* is functional with regenerative properties. Further it is stated that the scaffold successfully has been transplanted to rats into an orthotopic position where it showed patency and normal function during a 2-week period.

Comments:
The structural evaluation of the transplants *in vitro* was mainly done by immunocytochemistry. These evaluations are hard to evaluate since characterisation of the various antibodies is lacking; no controls concerning specificity or cross reactivity are reported. The mechanistic studies on the transplant involve strain, force and distensibility. These studies are of interest in order to characterize functionality of the transplant but even more interesting to explore is peristalsis. The latter has, however, not been examined. A muscular layer is reported to have been developed (fig 6) and since peristalsis is of utmost importance for a functional oesophagus possible motor activity should have been evaluted.

The evaluation of the orthotopic transplantations is difficult. The same hurdles concerning antibody specificity as in the *in vitro* studies exist. What do the antibodies bind to? The quality of the micrographs is mediocre, no detailed cellular localisation of the antigen-antibody complexes can be determined. In particular, neovascularization and reinnervation should be better displayed. Fig 7h is stated to illustrate ”the whole graft (marked with dashed lines)”. This micrograph is puzzling, in the middle of the marked transplant a striking thickening is seen – resembling healing after an anastomos - while in the area marked with “dashed lines”, where the anastomosis are stated to be, the tissue is seamless and in continuity. The scale bar indicates 500 μm which reveals that the transplant is only 4 mm long. A certain degree of shrinking probably occurs due to preparation but original length of the transplant is reported 15 mm, indicating a notable shrinkage. Fig 7 is entitled ”fourteen-day orthotopic transplantation” but fig 7a shows a scaffold with GFP-positive cells after 3 weeks *in vitro*. Thus, this figure shows a transplant before transplantation. Single cells are hard to identify and what information this figure adds is not obvious. Highly notable is the information in fig 7c. This shows that the animals, sham as well as transplanted, have a severe weight loss indicating that the animals cannot feed and are severely malnourished. It is also notable that nowhere in the article is the number of animals used in the different groups indicated. The only information given is that there are 10 animals in the group.

Notification of suspected scientific misconduct

OE Simonsson (OS), M Corbascio and K-H Grinnemo reported to KI Vice-Chancellor Prof. A Hamsten on 2014-06-24 that the article contains ”inconsistences, misrepresentations and examples of scientific misconduct”. The notifiers confided 4 points that they considered to mislead the reader into believing that the esophageal tranplants are functional with regenerative properties:

Point 1. Fig 8a, a CT scan illustration. The notifiers state that the information that the rat in question has a probe inserted into its esophagus for delivery of contrast by the time of the examination has ben left out in the article. This leads readers to misinterpret the image and to believe that the probe represents the inner esophagus lumen which in turn gives an incorrect perception of a ”smooth internal surface”, a property described in the figure legend.
The performer of the CT examination is one of the notifiers (OS) and additional images to prove the above statement are attached.

Point 2. The notifiers also object to that the finding that transplanted animals lost less body weight than shamoperated (fig 7c), by the authors are used as an argument to draw the conclusion that this is the first successful development and transplantation of a decellularised biological esophagus scaffold that after repopulation of allogenic stroma cells has generated a functional transplant with regenerative properties *in vivo*.

Point 3. OS informs that he has performed a CT scan examination of a transplanted rat 16 days post operatively –not mentioned in the article. In these images hair and food were noted in proximal (upper) esophagus which, according to the notifiers, indicates dysfunctionality of the transplant. The notifiers also claim existence of CT scan evidens for that the animals suffered aspiration pneumonia when gavage feeding ended and they were fed food with soft texture.

Point 4. The notifiers further question how so few GFP-positive cells before transplantation (illustrated in fig 7a) can correspond to the image (fig 7b, showing the transplant *in situ* photographed with a dissection microscope) after transplantation, and they suggest that, in fig 7b, images from sham animals have been confused with transplanted animals.

Comments PM 2014-08-03

PM rejects all accusations. Begins by describing criteria for the use of scaffolds and necessary properties of grafts and states that the, here described, esophageal transplant meets all these requirements and that the complainants at no points can argue that the transplant is dysfunctional or that it not has regenerative properties.

The four points of allegation are discussed:

Point 1. PM claims that OS, who performed the examination, never conveyed to any of the other authors that the probe for contrast administration passed the transplant. Nobody from PM research group (ACTREM) had access to the CT scanner, OS performed the CT examination and provided Sebastian Sjöqvist (SS) with the illustration used for publication. Three sets of data were generated – one without contrast, one with contrast but with the probe proximally in the oesophagus and one with contrast but without any probe. None of these showed any probe through the transplant. (*My remark: This is unpublished images, postoperative day is unknown and since no probe has passed the transplant, according to PM, it is not surprising that no probe ca be identified in that region. How "contrast without probe" was performed is unclear. It is also unclear on which animals these examinations were performed; no other documents mention these investigations further).*

The conclusions that the organ was "patent" and that the inner surface is smooth (i.e. epithelialized) were by the authors only partly drawn from the published image (fig 8a). Other images showing the same results is claimed to be at place but this particular image (fig 8a) was best suited for publication. Of importance to note, according to PM, is that patency of the organ is also illustrated by fig 2a-d. That the inner surface is smooth is also shown in fig 2a-d, reepithelisation is illustrated in fig 2e-f and by scanning elektronmicroscopy, fig. 2g-h. PM also claims that the diameter of the probe is less than that of the esophagus and that it consequently is unable to fill the lumen. PM also highlights that a static oseophagogram on a living animal gives a homogenous diameter of the organ.

Several key findings are listed in order to emphasis that the CT scan picture was of no central importance to draw the conclusions in the article.

Point 2. That the differences in the weigth curves sham *vs* transplanted animals formed the basis to conclude that this is the the first successful development and transplantation of a decellularised biological oesophagus scaffold that, after repopulation of allogenic stroma cells, generated a functional transplant with regenerative properties *in vivo* is dismissed with the argument that all, in the article presented, data formed the basis for this conclusion (refers to last paragraph in Discussion). PM underlines that transplanted animals have ”more favourable weight curves” which is suggested to be due to their more patent esophagi. He also points out that the animals consumed food and that the tranplanted animals only lost 12% of their body weight on day 7 postoperatively compared with sham animals that lost 25%.

Point 3. The claim that hair and other materials in the esophagus indicate dysfunctionality is met with that this is a normal finding, also occurring in native rat esofagus, and does not indicate dysfunction.

Any aspiration pneumonia was, according to PM, if it ever existed, clinically asymptomatic. The animals were reported to ingest food, groom themselves and they showed no signs of nausea or active pneumonia. PM also points out that no evidence of aspiration pneumonia has been reported from the CT studies performed by OS during the study. PM remarks that the CT examination referred to was performed on an animal 16 days post-operatively initially during the study when multiple complications arose due to lack of experience and that it can not be considered representative. Adequate post-operative treatment was, after the initial trial period, discussed and worked out in consultation with the veterinary and the regional animal ethics committee.

That the animals received nutritional support via gavage is completely rejected.

The notifiers speculation that postoperative follow-up was only 2 weeks due to the animals' poor health status is dismissed; two weeks was the planned end point of these experiments.

Point 4. The notifiers’ suggestion that a confusion of transplanted and sham- operated rats may have occurred in Figure 7b is dismissed as unprofessional; the implant is transparent and reddish due to the medium that it has been immersed in.

PM further points out that a conflict between the notifiers and PMs research group exist. He also expresses great surprise that OS withdraw his co-authorship when proofs of the article arrived. Any suspicions of scientific misconduct think PM should have been communicated to the other co-authors earlier during the work and at the latest when submitting the article for publication.

Sjöqvist lab books (11 files)
Described below only with the period it covers + keywords of main content. Ethics permits noted: diary no S175-09, S131-11, S160-11.

File 1. Aug 2011-Feb 2012, begins commendably with KI guidelines for planning, conducting and documenting experimental research.
Contains practical information, biopsies trachea, heart, oesophagus, pig, rat, isolation of stem cells, histology, immunocytochemistry. Decellularisation of esophagus in Jan 2012.
24/1, 2012 Lab animal sci class

File 2. Feb 2012-April 2012, histology, immunocytochemistry, *in vitro* culture, decell, oesophagus, scanning electronmicroskopy protocol, proteomics

File 3. April 2012-Sept 2012, Proteomics, ABEL kit, cell activation, neutrophil activation

File 4. Sept 2012-March 2013, Immunocytochemistry, native oesophagus, electric stim, reseeding

File 5. Mars 2013-May 2013, reseeding, stress tests, burst pressure. The first orthotrpl noted 23/4 (6 mm reseeded esophagus trpl to cervical esophagus), 24/4 the rat died. In this file a number of worring comments appears: "sham rat dead in cage", "sacrifice ortho 4, due to excessive weight loss and ill health", "sacrifice sham- 3 days survival", "mucos in mouth", "orbital thightening", "rat died from eating (paper?) – got caught in esophagus", "not still well".

File 6. Jun 2013-Sept 2013, continuing transplantation experimentation. CT scan, notable comments: "very poor health status", "dead in cage","dyspnea?", "very stenotic", "ortho 6 to CT", "chewing air", "dead in cage", "metabolic cage for 14 days"

File 7. Sept 2013 – Oct 2013, TX of 7 mm seeded scaffold. Fistulations are noted. Sacrifice: pilo, back, tenting, orb is reported and "graft ruptur"

File 8. Oct 2013 – Nov 2013, Immunocytochemistry, calcium studies

File 9. Nov 2013 – Dec 2013, IHC of "ortho", CT,

File 10. Jan 2014 –Feb 2014, Biomecanics, IHC

File 11. Feb 2014 –July2014, IHC, EM, subcutaneous implantation, misc projects e.g. trachea

Review by Professor emeritus Bengt Gerdin 2015-05-13

Bengt Gerdin (BG) has, on KI's assignment been the expert examining the allegations concerning suspected scientific misconduct.
In assessing whether scientific misconduct has taken place the investigator has assumed: 1. The European Code of Conduct for Research Integrity, www.esf.org; www.allea.org, 2.God forskningssed; Vetenskapsrådets rapportserie 1:2011, 3. Lag (2003:460) om etikprövning av forskning som avser människor (etikprövningslagen), 4. http://codex.vr.se och 5. The Committee on Publication Ethics (COPE), http://publicationethics.org.

The review first addresses the four points in the notification on suspected scientific misconduct:

Point 1. BG asks himself: what morphological support exists to claim that the graft is covered by epithelium? The scant method description is highlighted but OS, on direct request, states that the study illustrated in Figure 8a is made five days postoperatively, the animal was alive when the examination started but died meanwhile, and that the probe most likely was a 18G feeding catheter from AgnTho with an outer diameter of 1.27 mm and an inner diameter of 0.838 mm. (*My remark: this information appears first in this document*).

BG has, from OS, received DICOM-files of the CT examination in question (performed 2013-04-30). Together with Dr P Liss, Akademiska sjukhusets röntgenavdelning, he re-examined the images. This re-examination reveals that the structure seen has the same diameter all the way from the oral cavity and down, and that it is the probe containing contrast medium. The conclusion is that the conture interpreted as ”a smooth and patent oesophagus” by the authors is not this at all and that the CT examination only shows that the transplant can be passed through by the probe used.

BG remarks that, based on information in "author contribution", none of the authors are in the position to take responsibility for the CT examination which, in his opinion, diverts from the Vancouver Declaration as well as Nature´s regulations. Two of the authors (Rainer Heuchel, RH och Ying Zhao, YZ) are stated to have “been involed” in the CT examinations. The examination was performed by OS (one of the notifiers) whom at the time of submission to Nature Communications was a co-author but gave up authorship at the time of proof reading. BG remarks that Nature does not require the approval of all co-authors, only that the "corresponding author" declares that "all the listed authors have agreed all of the contents".

BG concludes that from the perspective of scientific misconduct, the key issue is not what the image shows or does not show , but whether the authors , deliberately or negligently, presented research results obtained by other researchers to support their main theory, even though none of the authors had knowledge of critical methodological or interpretation related details. BG deems that this is the case.

Point 2. Concerning the animals’ body weights. Do they have a functional esophagus? The statement in the article that ”All animals survive the 14-day study period, with patent and functional grafts, and gain significantly more weight than sham-operated animals” (Abstract) is questioned since no ”weight gain” what so ever could be noted in any of the groups. BG again comments on that the article has scant methodology descriptions. After having contacted ”the institution” BG finds that the weight curves (fig 7c) is based on 2+2 animals who lost 124, 128, 170 and 191 gram, respectively, during the postoperative period. This indicates that food intake was not sufficient and BG further emphasizes that the figures should be compared to a normal adult rat which gain approximately 50 grams during a 2 weeks period. If the operated animals have had functioning esophagi then their weight loss should have ceased, eventually stabilized and towards the end of the period the rats should again gain weight. BG finds that the reported significant difference in weight gain between the groups can not be reproduced, which is not considered unexpected since n = 2 in each group.

In summary, the data reported in the article does not support the authors' interpretations. BG does not consider this is a so great deviation from good scientific practice that it constitutes scientific misconduct but should have been handled in the peer review before publication.

Point 3. CT scan images showing ”food and hair mixed with contrast fluid lodged in the proximal part of the reseeded esophagus” have from the notifiers been clamied to exist but not published. BG finds that this single investigation can not be considered representative for the entire group.

Point 4. Based on present material BG can not judge if and to what extent fig 7a actually shows GFP-positive cells growing on the scaffold.

Other comments from BG:
The laboratory books show that many animals died after transplantation which means that the statements in the article expressing that "all animals survive the 14 -day study period, with patent and functional grafts" are misleading.

Fig 7 is claimed to show desmin-positive cells and angiogenesis, BG think it rather looks like pericytes which appear earlier than arteriols.

A typo error is noted (0.09 ska vara 0.9).

How ”patency” was calculated, how the measurements were made, how an accuracy of 0.01 % can be achieved and how "patency" can be expressed as the "ratio between the diameter of the graft and adjacent native esophagus" are questioned.

The authors claim that a postoperative period of 14 days in the rat corresponds to a period of one year in humans is questioned. None of the above is, however, deemed to be such a deviation from good scientific practice to constitute scientific misconduct.

The lab journals and records made available are by BG considered deficient. It is impossible to follow the experimentation from the records to the article and vice versa. It is however, not considered such a deviation from good scientific practice to constitute scientific misconduct.

Overall comments from BG:
At least one significant deviation from accepted practices in research and a number of basic scientific flaws exist in the work. Deviation from accepted practices in research is the use of research findings from other researchers, whose quality and origin none of the authors can take responsibility for. Probably due to lack of insight into the methodological details, have the results been presented incorrectly thus giving unfounded support to the main thesis of the article. This is at least reckless. The main author is responsible for this, which constitute scientific misconduct.

The scientific shortcomings highlighted by the notifiers and himself are matters that should have been handled by the journal's peer review system. The shortcomings are, however, of such a nature that they may give rise to misinterpretation of the results presented.

Comments from PM and co-authors, June 2015

PM begins by rejecting all accusations and announces the presentation of new evidence to show that the allegations are totally unfounded.

Focuses thereafter mainly on the point that BG assessed as ”scientific misconduct " - to present results of which the authors can not fully take responsibility for. As "corresponding

author" PM was underlined as responsible. PM argues that: SS and JH joined and were instructed by OS during the CT examination. OS has been involved in the writing of the article in particular concerning CT scanning, methodology and interpretation and has approved the manuscript. PM dismisses the accusation of scientific misconduct and expresses that he take on the responsibility for the interpretation of the CT scan examination, although he did not perform it.

SS points out that there is a more recent definition of the term "scientific misconduct" than used by BG, which he assumes. SS also reports that there are past and ongoing conflicts between PMs research group and the notifiers and that he personally has been exposed. Regarding BGs report he noted that the notifiers were interviewed but that he, for example, has not been invited to any. He complains that BG talked to media before those accused had received any information about the case. He also thinks that BG has misinterpreted the Nature Publishing Group's guidelines regarding the responsibility of "corresponding author" in a "multi- group collaboration". In response to BGs comments on the four allegation points notified he states:

Point 1. A CT scan can never show if the transplant is reepithelialised or not. Not until this examination is combined and supported by other examinations the research group can assure that this is the case.

OS performed the CT scan examination but SS and JH were introduced to the technique by OS. Further OS was active as a writer until he by mail (and phone) withdrew from author list due to "no scientific contribution" in connection with arrival of proofs. SS maintains that the published CT scan image shows a "patent esophagus with a smooth internal surface ". Furthermore, he emphasizes that the CT image was not central in drawing the final conclusions.

Point 2. The animals' negative weight curves were due to the surgery and not the esophageal transplant. BGs difficulties to reproduce the statistical difference between the two groups can SS not explain or understand. The fact that there are only two animals in each group is explained by that it was the referee who wanted the data to be presented as weight/day; more animals could have been included but these were not weighed on a daily base.

BGs comments that the animals never"gain weight" and that their age was 8 weeks are recognized as incorrect.

Point 3. The CT image showing hair and food in the esophagus was generated from one single animal 16 days post surgery and therefore not representative.

Point 4. SS refers to PMs previous answer (2014-08-03) to this notification.

Other comments from SS:
The statement in the article that "all animals survived" was considered wrong by BG but is defended by SS as a "glitch" since the animals died due to "scaffold dissection".

The question whether desmin-immunoreactivity in the preparations represents pericytes or arterioles think SS can be left open but believes that it can not form the basis of "scientific misconduct".

Esophageal "patency" is not difficult to estimate, according to SS; the tube is opened longitudinally, measurements are made and expressed as the ratio between the graft diameter and adjacent esophagus.

BGs comment on that the laboratory records were incomplete and messy is met with the statement that they were detailed enough to reveal evidence behind the study data and conclusions. Health parameters such as respiration, general health, locomotion, etc. were documented daily.

Finally SS underlines that the method needs further evaluation and development, primarily it needs to be tested in a large animal model, with complete transplantation of the entire esophagus

Mei Ling Lim: is a stem cell and molocular biologist claiming that the CT scan image shows ”a smooth and patent oesophagus” and that the weight curves reflect surgery and not the transplant. Attests that the article does not contain any fakes, plagiarism, omission of data or deceptive results.

Greg Lemon: calculated the number of cells necessary for seeding the modified tissue grafts. Consider the notification as a revenge attack on ACTREM.

Karolina Kublickiene: performed "distensibility testing" and argues against any scientific misconduct; the team showed good teamwork and high passion.

Ying Zhao performed micro CT scans. Add a list of examinations performed, of which one is presented in the article (fig 2 n and o; esophagus scaffold and esophagus *ex vivo*)

Miguel Burguillos: evaluated inflammatory response using immunocytochemistry and has no comments concerning the CT scans.

Peter Damberg: generated MRI data from ex-vivo preparations. States that his efforts and expertise are limited to MRI.

Rainer Heuchel: has, from OS, never heard any objections or accusations during manuscript preparation.

Johannes Haag: states that the allegations came unexpectedly and after conflicts between the two teams ACTREM and "the heart- group". JH is surprised that only one reviewer investigated the allegations and that interviews with the notifiers, but not with others took place. JH further expresses surprise that media was contacted before the Vice Chancellor's decision and before those accused were informed. JH participated in CT imaging, and states that OS never raised any doubts towards these studies. Writes that OS instructed himself and SS, later YZ and RH performed CT examinations. Claims that CT imaging was a very small part of the study and that it did not play any decisive role in the main hypothesis of the work.

Christian Ibarra: only involved in studies on ”calcium signaling” and cannot speak for the other parts of the work.

Philipp Jungebluth: is SS co-supervisor and has closely followed the whole process and certifies that generation, interpretation and presentation of data are correct; nothing has been

fabricated or modified. PJ think that OS has misled the other co-authors since he did not express any objections concerning the CT image or the investigation earlier.

Bertrand Joseph: performed the macrophage assays (Figure 5). He underlines that the CT scan interpretation never include the word "epithelium" and that the CT image is an illustration, if removed it would not affect the main conclusions in the article. Believes that OS should have been mentioned in "acknowledgments" but also that OS should, in turn, have announced his suspicions of scientific misconduct earlier in the process.

KI decision 2015-08-28

KI does not see any conflict of interest for the reviewer BG.

KI notes that the authors' comments on the external investigation have brought new material that has been decisive for the assessment of the notification.

KI shares, in principle, BG´s assessment that PM is responsible for that it among the authors exist the skills to properly present and interpret all data included in the article. The comments from PM and co-authors have credibly challenged BG's assessment that none of the co-authors participated in or had knowledge of the CT scan methodology. It is noted that OS left the writing process after the manuscript has been accepted and was in proof.

KI notes that the transplanted animals' weight development was "just as good" as that of sham operated animals.

Decision:
PM has acted carelessly, but is not guilty of scientific misconduct.

Decided measures:
Determines that certain conditions have emerged regarding PMs research and that this does not meet the highest standards of quality.

PM should ensure that an erratum is submitted to the journal based on the observations that have emerged in the investigation.

A meeting with the Vice-Chancellor, the Head of Dept of Clinical Science, Intervention and Technology (CLINTEC), Head of the Ear, Nose and Throat Diseases and PM where current circumstances are reviewed is to be held.

The head of CLINTEC should consider whether this decision will have consequences for PM and co-authors.

The decision is signed by the Vice-Chancellor Professor Anders Hamsten

BGs close-up comments 2015-12-21

NOTE this document apply jointly to two reported notifications; the current and a patient investigation. BG states that when he compares his request for extradition of documents with what was actually handed over, and with what, in retrospect, was mentioned in various comments he realizes that significant amounts of original material was never extradited. PMs

concealment of material is considered remarkable and is considered to have sabotaged the investigation. The consequence was that "evidence" supporting the respondents arrived in retrospect, they were never scrutinized and in addition they constituted the basis for KI's decision.

Specific comments regarding the article: Based on PMs own statement (2014-08-03) that no one from ACTREM had access to the CT scanner or software or the expertise to make any 3D imaging and therefore dependent on OS to perform these studies is it still BG´s view that the research team could not take responsibility for the results included in the article and his conclusion on research misconduct persists.

Other general and specific comments on flaws in the investigation concern the patient investigation.

Correspondens PM and Nature Communications

2015-08-31, PM informs the senior editor Christoph Schmitt on the KI Vice-Chancellor´s decision that scientific misconduct has not occurred in the current article but that certain audit of the article is needed.

2015-10-19, editor Niki Scaplehorn respond that they will handle the potential errors in the article with a corrigendum, which will be peer reviewed before publication. Regarding the CT scan (Fig 8a), the journal will consult a technical expert to investigate whether the image shows a "patent" esophagus and not just the probe. The journal also requests further clarification regarding postoperative time of scanning, detailed description of the method as well as which and how many animals were scanned, number of animals included in Figure 7c, and some further corrections mainly of typo character.

Report County Administrative Board, 2016-03-09

The County Administrative Board, Stockholm, reports after inspection according to the Sw. Animal Welfare Act § 24, on PMs animal activity in the county. Eight different ethics permits (+ supplements) to perform animal experimentation have been held by PM. The control is administrative of the type directed, planned and without announcement.

Regarding the article Sjöqvist et al, Nature Communications 2014, 15;5:3562, the given animal ethics permits with diary numbers S12-13 and S36-13, as well as the addendums S23-14 and S160-11 are found applicable.

The County Administrative Board comments that the article does not state that the animals received analgesia. For animals with subcutaneous grafts it is noted that the regional animal ethics committee had approved this to be performed without pain relief (S160-11). It is also noted that animals that have undergone orthotopic transplantation were not fed according to the ethics permit S12-13 but that a later addendum changed the feeding regime. Otherwise, the County Adminsitrative Board deemed the description of all procedures and measures to comply with what is stated in S12-13.

The Approved Manager at the animal facility in Huddinge, Moustapha Hassan, submitted checklists for anesthesia of 10 rats (surgery June -November 2013) and photographed

laboratory records of 11 rats operated from March to May, year: unknown. (*My remark: please NOTE these records describe rats operated with trachea transplant, not esophageal)* The County Administrative Board finds no reason to distrust the material and expresses no critique on animal pain relief. They comment on that 10 rats after transplantation of the esophagus were not fed according to the description of the ethics permit, but noted that this deficiency is remedied and leave the case without further action.

”Questions asked and data requested concerning O1-2016”

The request was provided in Swedish and English; PM´s reply was in English. Below is a summary.

Question 1a: The county administrative board states in their report (2016-03-09) that the article Sjöqvist et al Nature Commun 2014, 15;5:3562, is based on animal ethics permits S12-13 (with 2 addendum) and S160-11. In total 94 rats were claimed used for *in vitro* studies and as donors, 3 rats were used for subcutaneous transplantation and 10 rats for orthotopic transplantation.
Is this correct? If not please state correct numbers.

Reply: Not quite correct. 94 rats in total: 9 subcutaneous implants, 67 donor animals/ *in vitro* studies and 16 received tissue engeneered grafts. Of the latter 4 underwent CT scanning.

Question 1b: please provide log records (försöksdjursjournaler) in accordance with the Swedish Animal Welfare Ordinance, L150, SJVFS 2015:38, chap 8, of all transplanted animals (subcutaneously and orthotopic; sham and transplanted) used within the above mentioned permits. In addition to the mandatory information, as stated in L150, also provide the daily assessment points (0-0.4) on the different parameters included in KI health assessment checklist and body weights (for each animal throughout the entire postoperative period).

Reply: Regrettably, the logs for all animals were lost when moving from one animal facility to another. The animals were checked on a daily basis and if any animal scored higher than specified endpoint, they were sacrificed. All animals were in good condition and no sacrifice was necessary (except for the animals that had dissection of the graft pre-implantation, G201 and G202). Refers to appendix 1 which lists weights and health parameters for 9 of the 16 animals: 2 sham and 7 orthotopic transplanted of which 2 were killed on postoperative day 8. (*My remark: health parameters were not scored in accordance with the KI's template but by using "+" or "-". If one assigns "+", the lowest possible KI score 0.1 so had the latter two animals a score of at least 0.8 by the time they were removed from the study. All animals reached score 0.4 already postop day 3. Humane end- point was> 0.3).*

Question 1c: which of the rats in *1b)* were included/excluded in Results in the article?

Reply: All animals were included in the results, with an exception of, G201 and G202.

Question 2: Fig 4 m, r, t, u, v and w: please provide the original data for these figures.

Reply: The original data is provided in appendix 2.

Question 3: Fig 5 g and h: please provide the original data for these two figures.

Reply: The original data is provided in the appendix 3.

Question 4: Fig 6 m, n and p: please provide the original data for these three figures.

Reply: The original data is provided in the appendix 4.

Question 5: Fig 7c: please provide the original data on which this figure is based.

Reply: The original data for figure 7c is provided in appendix 5.

Question 6a: Documents provided from KI concerning the previous investigation of scientific misconduct reveal that Fig 7c. is based on n=2 in each group. Is this correct?
Question 6b: If so, how do you justify expressing the data as mean ± SD and calculating statistical significance between two such small groups?

Reply: a) Yes, it is correct. b) In agreement with Nature Communications an *Erratum* with an updated graph including all individual data points will be presented. A reviewer suggested the "daily weights" during the peer-reviewing by Nature Communications. Weights from several time points from several animals were at hand, but no *daily* weights. Hence, Fig 7c was presented. More replicates would have been favorable, but time restriction prevented us from doing more experiments. It was a mistake to present the graph as mean ± SD, instead each point should have been plotted.

Question 7a: According to the above mentioned documents from KI it has also been stated that Fig 8a illustrates a CT scan performed on an orthotopic grafted rat 5 days postop using a feeding probe through the esophagus to administrate the contrast at the level of manubrium sterni. The rat was alive when the examination started but died during the procedure. Correct?

Reply: Yes. Regarding this point, we are still awaiting results from Nature Communications' ongoing peer-review of the CT scans. In our opinion, the responses prove that the description of the CT image in the publication is correct. Please see the attached documents (appendix 6 and 7), in particular Fig 3c, which shows protrusions on the structure which never could exist if the contrast was exclusively limited to the lumen of the probe.

Question 7b: Fig 8a is the only published CT scan image but the documents from KI indicate that additional rats have been examined by CT scanning. How many and by whom? At what time points postoperatively and by what procedures (contrast administrations, probes, etc)? Were the animals sacrificed in conjunction with the examination?

Reply: The animal was anaesthetised with Domitor/Dormicum and Fentanyl, placed in a supine position, a contrast delivery probe was introduced through the mouth and CT images were captured. In total 4 animals were CT scanned – at days 5, 7, 10 and 12. Each animal was euthanised after the CT procedure. The CT from day 5 is the published picture and it was done by OS with assistance by SS and JH, the others were done by YZ assisted by SS.

Qustion 8: The supplementary movie shows grafted rats. At what age, postoperatively?

Reply: they are around 12 weeks old. *(My remark: PM has not answered the question)*

Assessment

Based on my assignment to judge if that revealed by the documents represents scientific misconduct, I want to start by addressing the four original points of allegation:

Allegation point 1
Figure 8a, CT scan image in which the notifiers suggest that the presence of a probe within the esophagus at the time of the examination leads to misinterpretation and incorrect perception of a "smooth internal surface". The performer of the CT examination is one of the notifiers (OS).

BG, along with X-ray expertise concludes, after re-reviewing the images that the study basically only shows that the graft can be passed through with the probe and that, from a fraud perspective, the key issue is not what the image shows or does not show but that the authors, consciously or negligently, presented results obtained by other researchers to support the main theory of the study. This despite the fact that none of the authors had knowledge of critical methodological or interpretation related details. The principal author is considered responsible for this, which constitute scientific misconduct.

I fully support the above conclusion. There are still uncertainties regarding who conducted and who was responsible for the CT scan examinations (see also below point 3). PM has declared that he takes responsibility for the image and the investigation, but none of the authors have been able to clearly demonstrate that they have the basic skills required, or even that they have access to the equipment. The notifiers' main objection is that the authors have failed to describe that the probe used to administer the contrast has been inserted through the transplant. One wonders then if the executor of the investigation (OS), in turn, has described this for his co-authors? What responsibility did OS take on concerning "his" investigation and how accompanying results were used when he decided to resign as co-author? Different reasons have been put forward to explain why OS ended his co-authorship at the time of proof reading, but whatever the reason; I presume that he has had the opportunity to design a proper method description. Has this been deleted after OS dropped out? I find no indications that this is the case.

Based on the main issue of possible scientific misconduct a new re-examination of the CT image/images is not likely to reveal new information. This especially since both the first and last author claims that the image in no way was crucial for the conclusions drawn in the article. In addition, it subsequently emerged that the published CT examination was performed on an animal 5 days after transplantation and can therefore not be considered representative for the entire group of transplanted animals, 14 days after surgery. However, the current CT examination is, in the article, described under the titel "Histological and functional graft evaluation postmortem." This section begins with: "All animals survived to the study endpoint of 14 days ..". Thus, the reader, falsely, consider Fig 8a as an illustration of an animal 14 days after orthotopic transplantation i.e. the reader is misled.

Furthermore, it is noteworthy that, for this investigation, no ethics permit to carry it out under the conditions stated by both OS and PM, namely that the animals received contrast via oral gavage, is at hand. As no details about how the CT scans were performed are found in the article's method description I assume that OS and PMs consistent versions are correct. The animal ethics permit (diary no. S12-13) authorizes so-called low dose animal-CT examination. The procedure for this is described as the animals are anesthetized with isoflourane and

placed in a box heated to 37 degrees for a maximum of 15 minutes. Nothing in the permit allows for contrast administration by oral gavage. This deviation from the permit is, in my mind, considered as scientific misconduct and also a violation of the Sw. Animal Welfare Act (SFS 1988: 534 §2 and §10). As PM assumes all responsibility for the CT examination and is assigned principal investigator on the ethics permit I conclude that this falls heavily on him. OS role as performer entails, however, that he is partly responsible for having deviated from the given animal ethics permit. According to PM also YZ performed 3 CT scan examinations on living transplanted animals using the above described methodology.

Allegation point 2

The notifiers claim that the finding that the transplanted animals lost less weight than those sham operated (Fig 7c) is used as an argument to conclude that this is the first successful development and transplantation of a decellularised esofagus scaffold that after repopulation of allogenic stromal cells can be used as a functional graft with regenerative properties *in vivo*. This is not correct, also the authors themselves argue against this. No such direct link is expressed in the article. However, this is not the most urgent reason for a closer investigation of the weight curves. BG complains, quite correct in my opinion, that if the operated animals had a functioning esophagus then their weight loss should have stalled, eventually stabilized and towards the end of the postoperative period the rats should again gain weight, that one can not speak of any "weight gain" for any of the two groups, and he also questions the demonstrated statistical significance between the two groups based solely on two individuals per group. However, BG does not believe that this is such a big deviation from good scientific practice to constitute scientific misconduct.

From my side, I want to emphasize that an animal which during a 14 day period loose approximately 40 % of its body weight suffers from excessive weight loss and has a severly reduced health condition. Spontaneously, I wonder how the regional ethics committee (Stockholm South) could authorize a procedure that involves such a large body weight decline. Consulting the animal ethics permit of orthotopic transplants (diary no. S12-13) reveals that the humane end-point for these experiments is described as "animals are observed and weighed daily and euthanized at signs of impaired health (> 0.3 at KI template or > 0.5 the first 24 hours)". KI template is a standardized template for the assessment of animal health for small rodents and is based on a "score" for various health parameters such as general condition, porphyria/eye inflammation, movement and body posture, piloerection, skin condition and weight. Each parameter is scored 0-0.4 according to specific criteria and is added together. If, as stipulated for the current experiments, the total score is > 0.3 (or > 0.5 the first 24 hours after surgery) then the animal should be removed from the study as it is then deemed exposed to "unnecessary suffering". If we only look at the body weights (all other parameters left unchecked) we note that already on day 2 postoperatively the animals reaches 0.4 points (because they exhibit more than 10 % weight loss, as compared to their weight before surgery). 40 % weight loss is not even included in the KI template; 20 % is the maximum stated and may only persist for maximum three days postoperatively. After that a veterinarian must be contacted.

To consistently and systematically not respect the approved humane end-point as described in the animal ethics permit is, in my view, scientific misconduct and also a violation of the Sw. Animal Welfare Act (SFS 1988: 534, § 2). In his reply (2016-05-16) PM appends protocol for postoperative follow-up of 9 out of the 16 orthotopically transplanted animals. All 9 reach the humane end-point i.e. more than 0.3 points already at postoperative day 3, mainly due to

weight loss, and should then have been taken out of the study. The assigned principal investigator, PM, bears ultimate responsibility for obeying the conditions stated in the ethics permit but I also question whether other employees, including animal house personnel, veterinary and the Approved manager have taken on the responsibility expected in their respective positions. Furthermore, I wonder why the County Administrative Board's animal welfare officers and veterinary on performance of an animal welfare control did not point out the anomaly regarding the experimental humane end-point.

In this context the strange phrasing in KI's decision (2015-08-28) stating that the transplanted animals' weight development was "as good as" that of the sham operated animals needs to be pointed out. The word "good" is completely out of place in this context.

Furthermore, it is my believe that the principal investigator (PM) has misled the regional animal ethics committee by describing the orthotopic transplants like this (S12-13, section 8): *"This is a medium- difficult procedure that has been described in several publications e.g. Grikscheit et al J Thoracic and Cardiovascular Surgery 126: 537-544, 2003 and Lopes et al Diseases of the Esophagus 19: 254-259, 2006. With our previous experience from transplanting the trachea of rats, pigs and humans, we feel confident to do this type of surgery. After transplantation, the animals will be checked daily for any swallowing difficulties, if any difficulties arise, the animals will be killed. /./.We expect a very low incidence of this complication (< 5%). "*

If one reads the cited references one can conclude that Grikscheit et al operate abdominal esophagus, not cervical as in the current project, and therefore this report is not comparable. Lopes et al perform a type of esophageal transplants surgically equivalent to PMs application but these authors report severe complications for the animals. The rats are described as moribund (approx. dying) with significant weight loss (about 26% weight loss at day 14 postoperatively), malnutrition, fistula formations in the anastomoses, absence of peristalsis and no normal swallowing reflexes. Thus the cited references do not, in any way, strengthen the investigators claim that the current orthotopic transplantations are expected to be of medium severity, and that they have good experience of such interventions. My assessment of this is that PM deliberately misled the licensing authority, Stockholm south animal ethics committee, which I consider scientific misconduct. In this context it is of interest to note that both of the above mentioned references in the article are quoted in "Introduction" to substantiate an assertion that earlier attempts to experimentally transplant esophagus have been unsuccessful and resulted in e.g. strictures and fistula formation (refs no. 14 and 15 in the article).

Deviations from the assigned animal ethics permit (S12-13) are noted on at least one more point. According to the permit (S12-13, paragraph 6) it is stated that "animals are observed and weighed daily". This has only been done for two transplanted and 2 sham operated animals, which the authors themselves testify (SS opinion in 2015 and PMs answer 2016-05-16) since they could not fully deliver the weight curves displaying the animals daily weights, as requested by the journal's referee. The result of the referees request became fig. 7c with 2 + 2 animals.

In the same permit (point 9) is analgesia with Temgesic prescribed. Based on the article's description of the surgery this has not been administered. This was mentioned also in the County Administrative Board's report but caused no note since the Approved manager of the animal facility submitted a checklist for anesthesia in 10 rats operated on in June-November

2013, and laboratory records for another 11 operated rats. *(My remark: The latter was subjected to tracheal not esophageal surgery).* As further "proof" that pain relief was given postoperatively also copies of purchase orders of Temgesic (May and Aug, 2013) and a declaration from Dr P Jungebluth testifying that all animals were treated in accordance with the animal ethics permit incl. postoperative pain relief with s.c Temgesic 0.02 mg/kg every 8-12 hours for 48 hours were submitted. With this I can only say that we really do not know whether or not, and to what extent, post-operative analgesia has been administered.

In order to fully ascertain the extent (e.g. how many animals were subjected to the various interventions?, how was post-operative care?, has a veterinarian inspected the animals?, when were the animals killed i.e. what was the actual humane end-point? and how was each individual animal estimated according to the KI template?) of the above identified deviations from the ethics approval the animal log records (försöksdjursjournalerna) are needed. Those are reported to have disappeared in connection with moving to a new facility. Still today I can not, despite the County Administrative Board's report, direct questions and requests for disclosure of records and minutes even say how many animals were operated on (sham and ortho, respectively), how the post-operative period was or how long each animal survived surgery. The article states that 10 of the total of 94 animals were subjected to "ortho/sham surgery" while PM in response to a written request states that of the 94 animals, 16 were transplanted and of these were 4 used for CT examination while 2 were excluded. Postoperative protocols of 9 transplanted animals have been extradited; of these were 2 animals sham-operated and 7 orthotopic transplanted. Two of the latter were taken out of the study on postoperative day 8, making it likely that the article is based on 2 controls (sham-opererated) and 5 orthotopically transplanted animals. Based on this, I question the overall reliability regarding the statistical calculations and also the author's conclusions drewn from the reported results (see also below "handling and presentation of data").

Allegation point 3

The notifiers claim additional CT scan evidence of aspiration pneumonia and poor health status.

BG does not believe that this single study can be considered as representative of the group. I agree but would, from the perspective of scientific misconduct, again call attention to (see above points 1 and 2) the identified deviations from the ethics permit when performing CT scan examinations and deviations from the approved humane end-point. Any suspected cases of poor health and suffering of the animals had been a non-issue if the approved humane end-point was respected.

Allegation point 4

The notifiers reported suspicion of confusion of images from sham and transplanted animals in Figure 7 is commented by BG. Based on the given information he is unable to assess to what extent Figure 7a shows GFP - positive cells growing on the transplant. Obviously I neither can tell whether any "confusion" has taken place or not. I have difficulties in understanding the notifiers' intention with this complaint. Comparing Figure 7a (GFP immunocytochemistry) with Figure 7b (an image taken with a dissection microscope during surgery) is, in my opinion, not relevant.

The substandard quality of Figure 7a connects, however, to several other scientific flaws in the article that caught the attention of BG. These were considered to be of that kind that they should have been handled by the journal's peer review system but BG states that they may give rise to misinterpretation of the results presented.

I definitely agree with this conclusion. Whether the article in its entirety is misleading or not is a more complex question to answer. The rationale and set up behind the current study is reasonable and innovative for the goal of developing a functional transplant to replace damaged/pathological esophagi. An interesting illustration of this is, for example, a new article in Lancet by Dua et al , in press , 2016 , doi: 10.1016/S0140-6736(15)01036-3. This article reports on the repair of damaged esophagus in a patient by means of a stent and tissue material from skin and muscle. Today, seven years after surgery the esophagus has healed and the patient has a functioning swallowing reflex. If the current article, Sjöqvist et al 2014, despite the indications of scientific misconduct herein identified, may contribute to the development of new functional grafts is entirely dependent on how well the results presented are based on reliable and reproducible raw data. Unfortunately, the authors do not convincingly show this to be the case. The methodology is sketchily described in the article. Original data can not be extracted from Sjöqvist´s scanned lab books, 11 files. These appear to be a mix of lab diary and calendar. They lack systematic protocols and summaries on, for example, which and how many analyzes were performed or how the analyzes were statistically handled and interpreted. Which and how many animals were included/excluded in the various experiments can not be determined. If this is due to deliberate research misconduct, negligence, carelessness or poor supervision can only be speculated.

Handling and presentation of data

Raw data has consistently been very scantly available even though BG already when preparing the first review of the article in 2014 requested the extradition of data for all trials. Availability and disclosure of data is a condition for publishing in Nature Group: "*A condition of publication in a Nature journal is that authors are required to make materials , data, code , and associated protocols promptly available to readers without undue qualifications"* and supported also by the National Archives regulations (RA-FS 1999: 1). When PMs response (2016-05-16) on the request for original data was received, data were obtained for Figs 4t and w , 5g - h, 6m, n and p and 7c while for Figures 4m, r, u , and v only means, spreads and number of observations (n) were disclosed. The following outlines the investigation made of these data. Generally, according to the article (Statistical Analysis, page 13), results are given as mean ± SD.

Fig 4m: shows protein quantification of extracellular proteins in native and decellularized esophagus. Results are presented as mean ± SD and seem to represent fig 4m in the article. Numbers (n) are stated in the article to be 3, but in the extradited material n = 8. This in itself does not change the graph's appearance or gives a more positive result, but indicates carelessness with data presentation. In legend to figure it is stated that "n = 3 biological replicates" which is not explained anywhere in the article. Statistical calculation is reported to be performed using ANOVA followed by Bonferroni post hoc test. Why was ANOVA used? The measurements made on concentrations of proteins are independent of each other. The Mann-Whitney U test had been an appropriate analysis since few data points are provided and the use of mean value is uncertain in such small materials. It had also been possible to see if the results were significantly different with a rank test but this can not be performed because raw data was not disclosed.

Fig 4r Shows how decellularized scaffold attracts blood vessels vs. collagen in the presence of phosphate buffer or growth factor VEGF, negative and positive control. Only means and spreads are extradited. N=3, and the results have been calculated using ANOVA (one or two way is not indicated) and Tukey's post hoc test, according to legend.

Fig 4t Original data submitted, consistent with published graph.

Fig 4u No original data available, only mean, SD and n=5. These are consistent with the published graph.

Fig 4v No original data available. Stipulated that mean ± SD is used but is illustrated with mean ± SEM in the published graph. This distorts the figure to a more positive interpretation as error bars become smaller and therefore variation within the group seems to be less. To present the results mean ± SEM must be actively selected in the Graph Pad Prism 5; SD is the standard setting.

Fig 4w This figure is described in the text (page 4, column 2 top paragraph) to show that "distensibility of the reseeded organ closely resembled the native" and in the figure legend as "lost distensibility after decellularization was regained after reseeding". These claims are an embellishment of data. The data have been calculated using ANOVA and Bonferroni post hoc test but no statistical significance levels are marked in the graph. Raw data has been submitted, n=3, but this does not match the legend where n=5. Repeated calculation shows that it is indeed possible to achieve a statistically significant difference in "distensibility" between reseeded and decellularised at lower pressures but at higher pressure this difference is equalized and no significant difference between reseeded and decellularized can be demonstrated, however, "native" significantly differs from the other two. This is true irrespective of if using Bonferroni or Tukey's posttest.

Fig 5g shows that macrophages are attracted to native transplanted esophagus while both decellularized and sham-operated showed lower infiltration rates. According to the article (page 13): "number presented is an average of macrophage cells per section". Furthermore, according to the article, n=3 animals per group and 8 sections per animal are cut and stained. On these 8 sections 3 photographs/section is documented for analysis. This gives at hand that the number of data points/group should be 3x8x3 = 72 observations/group alternatively 3x8 (if one calculates a mean value for each section) = 24 observations/group alternatively 3 (if one calculates a mean value of each animal). Data extradited gives at hand 38 observations in the "native" group, 28 in the "decellularized" group and 18 in the "sham" group. Either the investigation has not been carried out in accordance to the reported method description or data have been excluded.

Why and how these results are presented as "number per section" is enigmatic. How to standardized the cut section area and how is section area related to the three photographs taken? It is also unclear which tissues that were analyzed (the graft or the environment)? Shamoperated animals do not receive any transplant but are reported in the results section and in the graph. The results are presented in the graph as mean ± SEM, even though the article states the use of mean ± SD, which means that the spreads of the groups appear smaller than they actually were.

Fig 5h shows DNA quantification, normalized by sample weight, in subcutaneous transplants of native and decellularized esophagus. Significantly lower amount of DNA was found in decellularized esophagus. Each group consisted of three animals (n=3, which is also stated in the article) extradited raw data shows that the average is based on n=5 observations, which is puzzling. Again mean ± SEM is used, instead of the specified mean ± SD.

Fig 6m illustrates metabolic activity of mesenchymal stromal cells (MSC) seeded on the outer (muscle) and inner (epithelial) surface of the scaffold. A statistically significant difference between the two is shown in the figure. I question the scientific value of this comparison; the interesting comparison would be towards equivalent layers of native esophagus. Values are very low, 490 nm where the color change of formazan is measured may very well represent the blank. Raw data has not been disclosed, only the values from the plate. With raw data both the blank and the negative controls would be obtained and evaluation could be performed. The method description is completely missing from the article which means that information about at what wavelength the measurements were made or how corrections for blank and media were performed are missing. The trials can not be replicated without this information.

Fig 6n qPCR showing that MSCs grown in cell culture or on outer (muscle), or inner (epithelial) surface of the scaffold retains expression of CD90 but vary in desmin (muscle marker) and SP1 (epitelial marker). According to the article, the number of observations is 3, the extradited data gives at hand that there are only two observations in each group; all normalized to 1.0 in the "culture plate". Mean ± SEM is reported instead of mean ± SD as indicated in "Statistical methods". ANOVA (one-way ANOVA according to extradited data) and Tukey's post test are reported used for statistical calculation and this gives significant differences for all three markers, according to the figure. After recalculation, these reported results cannot be repeated. Only if using two-way ANOVA could statistical significance be achieved for desmin and SP1, but still not for CD90.

Irrespective of if significant differences are detected or not I strongly question how estimating statistical significance in groups with only two observations can be justified. Further, there is a discrepancy between the marked identity of the bars (muscle and epithelial) and the values in the data array. This results in that the graph shows opposite information as the figure legend. According to the graph desmin is missing in muscle-MSC while it in the MSC-epithelium is present in higher amount than in cultured cells. A corresponding contradiction exists regarding SP1.

Fig 6p illustrates the statistically significant difference in electric resistance for decellularized and reseeded scaffold. N=3 according to the article; original data report 7 data points/group, which is puzzling. Again mean ± SEM is reported, rather than mean ± SD. For the statistical calculations it is stated that a paired t test have been used. Why? I can not see that the measurements or the experimental set-up generate and require paired observations. Recalculation of data shows that if unpaired t-test is used, which should be the case when there is no mutual coupling between the relations, no statistical significant difference was obtained (p=0.48).

Fig 7c shows the animals weight curves after orthotopic transplantation and sham-operation, n=2. Statistical significance was achieved if normalized data were analyzed but not if the raw data was used (ANOVA followed by Bonferroni post-test).

Is the article's main message rooted in the results presented?

The main message of the article is that the authors successfully have developed a structurally and functionally decellularized scaffold repopulated by stem cells that are immunologically inactive. Further they then replaced esophagus in rats, with this transplant and they claim that these transplantations have been successful.

To argue that the transplant is functional, I would claim that a minimum requirement should be that transplanted animals eat and recover. This requirement is not met. No results in the study point even passably in this direction.

A number of statements in the article are not in accord with the actual results:

1) Abstract: "All animals survive the 14-day study period, with patent and functional grafts, and gain significantly more weight than sham-operated."
2) Introduction, last paragraph: "we successfully develop and characterize an oesophageal scaffold and subsequently demonstrate its functional *in vivo* regenerative properties in an orthotopic rat model".
3) Fig. 1, last sentence: "All animals survived until the 2-week time point, and further evaluations of the explanted grafts were performed."
4) Page 6, column 2, lines 3-9: "The animals, although initially immobile, quickly recovered and did not show any significant signs of pain or health impairments according to a pain assessment scale and general health scale. Interestingly, the transplanted animals' weight curves were significantly better than those of the sham surgery group from day 4 onward."
5) Page 6, column 2, lines 10-13: "All animals survived the study end point of 14 days and visual inspection showed patency of all oesophageal grafts."
6) Discussion, page 11, column 1, lines 3-6 from bottom: Notably, we successfully transplanted the reseeded oesophageal scaffold into an orthotopic position and showed patency and normal function over the entire 2-week study period."
7) Discussion, page 11, column 2, lines 2-3: "Interestingly, we found that the animals receiving the grafts had more favorable weight curves, a finding that could be attributed to the patency resulting from a circumferential segmental replacement of a functional scaffold…".
8) Discussion, page 11, column 2, lines 21-23: "In conclusion, we have developed a fast and reproducible perfusion protocol for decellularization of the rat oesophagus, resulting in an ideal biological scaffold with all necessary requirements for a transplantable graft."
9) Discussion, page 11, column 2, lines 5-8 from bottom: This is the first report of successful development and transplantation of a decellularized donor oesophagus, reseeded with allogenic adult stromal cells that yield a functional graft *in vivo* with regenerative properties.

The study's only limitation is described to be that it is not directly transferable to humans (Discussion, page 11, second column, second paragraph). This is suggested due to two reasons: 1) the fact that rat esophagus consistently has skeletal muscle and keratinized epithelium while humans have skeletal muscles only in the upper part and an unkeratinized epithelium, and 2) that the current experimental transplant was short and therefore probably not sufficient to replace human esophagus in various clinical situations.

Summary assessment

A number of weaknesses and negligence in data presentation and interpretation of results were identified. Selected statistical methods should be justified in "statistical methods" in the article; why and when 1-way, 2-way ANOVA and paired or unpaired analysis were used is unclear and seemingly random. The individual negligence in data management is of different magnitude for the article's credibility but taken together they indicate gross negligence.

The morphological part of the article is dominated by immunohistochemistry and illustrated with low-quality images. No controls concerning antibody specificity have been made (see also comments under "article"). It is possible that fig 9c is an attempt in this direction but unfortunately it is of little value to convincingly show that the authors have control on what molecules the individual antibodies show.

Taken together I think that the authors mislead readers to believe that these experiments were successful and that the results presented represent a significant step towards clinical use of tissue reconstructed esophagus.

The cumulative number of negligence stated in the above is breathtaking, given that the work is published in Nature Communications, a highly ranked international journal.

In summary, I find that PM is guilty of scientific misconduct due to:

- Refusal/inability to disclose all requested raw data for the results presented in the article.
- Refusal/inability to disclose animal log records (försöksdjurjournaler).
- Misleadingly presented, interpreted and described the results.
- Severely deviated from the animals ethics permit.
- Deceiving the licensing authority; the regional animal ethics committee.

PM is, in my opinion, also guilty of violating the Animal Welfare Act (SFS 1988: 534, §2 and §10) since he on several points has deviated from the animal ethics permit (neglecting the approved humane end-point, postoperative supervision and the procedures for CT scan).

Applied definition of scientific misconduct

The above conclusions on scientific misconduct are based on The European Code of Conduct for Research Integrity, www.esf.org; www.allea.org. This has guided my judgment rather than the Sw. Research Council´s, more narrow assessment. This because I believe that an experimental study in basic science published in an internationally highly ranked journal should be assessed from an international perspective. Furthermore, several of the authors have non-Swedish affiliations which further emphasize an international perspective. Desirable had obviously been an international "Code", but at present there is no such to refer to. My assessment of scientific misconduct had, with one exception, however, not differed if I instead used the definition given in the Sw. Research Council's ”God forskningssed”. The exception is the point which describes deviation from the animal ethics permit. According to the Sw. Research Council's narrow definition, these transgressions are to be handled in other contexts since they are not considered to have a direct impact on the scientific work.

I mean that PM has violated all the principles stated in "The Code". These principles include honesty, reliability, objectivity, transparency, responsibility for animals and human beings

incl. responsibility for future generations of researchers. Moreover, fraud committed through falsification and manipulation of data, plagiarism by not giving due recognition to OS CT scans, and failure to adhere to ethics permit. Regarding the definition of scientific misconduct by the Sw. Research Council, it is also fully applicable, it reads: scientific misconduct involves acts or omissions in connection with the research, which - consciously or by negligence- leads to false or distorted results or provide misleading information about an individual's contributions to the research (vetenskaplig oredlighet innebär handlingar eller underlåtelser i samband med forskning, vilka - medvetet eller av oaktsamhet- leder till falska eller förvrängda resultat eller ger vilseledande uppgifter om en persons insatser i forskningen).

Co-authors role in the article

The assignment also included to review the participating authors responsibilty for the content included in the article before and after publication. Based on my investigation, it is obvious from the article as well as the other documents that PM is the principal author. But contributing authors' role(s) are also to be discussed, at least that of some of the key persons. The article is published in Nature Communications. This journal's policy for authors determines that it is the "corresponding author", PM in this case, which is responsible for obtaining approval for the contents of the article from all authors. Nothing in the documents indicates that PM has failed to do so which ultimately therefore may be considered that all co-authors approved the article as published. The individual authors' efforts in this article are specified in "author contribution". The main reason behind these statements is that each contributing author's individual contribution is made clearly visible. By extension, this implies that the specified authors were also actively involved in the processing and presentation of data, or at least approved the way in which these were presented in the article. On this basis, should the individual co-authors role(s) be relative clear and linked with their individual contributions. In this case, however, it is uncertain since "author contribution" is surprisingly scarce. It contains several sweeping statements, for example is it indicated that MLL "Performed cell and molecular biology analysis". Does this include responsibility for the immunocytochemistry? Or is this methodology included in "pathological report" conducted by HK? Immunocytochemistry constitutes a fundamental part of the study, but none of the authors reported in "author contribution" is indicated to have performed it.

To individually review and evaluate the responsibility for PMs 22 co-authors regarding the article or the different parts therof is thus almost impossible to do. I can only emphasize and refer to the general responsibilities assumed for co-authors; approval of the publication and of the content of the article as well as their general responsibilities due to their employments (which may vary according to employment and affiliation). Despite this uncertainty a few key people emerges, which are described below.

Sebastian Sjöqvist

SS is first author and a PhD student. He is stated in "author contribution" to be the one who, together with PJ, operated animals and he was responsible for data collection. Implicitly I understand that he also was responsible for the data analysis and interpretation, as PM, PJ and MLL only helped in this task. Furthermore, SS, along with PM, wrote the manuscript. Proper documentation of experiments and their results and analysis falls heavily on the supervisor. Regarding the animal experimental work I assume that SS, as well as PJ, have attended mandatory training in lab animal science. PM is assigned principal investigator on the animal ethics permit meaning that it is ultimately PMs responsibility that the experimental animal activities are conducted in accordance with the permit. Overall, it seems as if SS, a PhD

student in the position of dependence with the supervisor PM, have been assigned a large (unreasonable large) responsibility for data collection, data processing and the experimental work load.

Philipp Jungebluth

From the documents I perceive PJs role as a research assistant in charge of the experimental work at ACTREM. PJ states (2015-06-24 , in the opinion of the co-authors, BGs examination) that he is SS co-supervisor and that he thouroughly has followed all efforts to generate, interpret and present scientific data and that he certifies that all data generated within ACTREM are properly presented and that nothing has been falsified or modified. PJ has also affirmed (2016-03-07, in reply to the County Administrative Board) that they by all means adhered to the animal ethics permit for each individual animal, including analgesia and humane end-points. Thus it is clear that PJ played a prominent role in the scientific work behind the article. Furthermore, I argue that PJ has actively misled by incorrectly certify both proper data handling and presentation of data, as well as proper handling of the animals from given animal ethics permit.

Oscar Simonsson

OS is MD and PhD student, affiliated to the Department of Laboratory Medicine, Karolinska Institutet, Stockholm. OS was co-author until the proofs arrived when he asked to leave the list of authors. This was announced in a neutral tone in a short email to PJ 2014-03-13, and the reason expressed was that he did not consider himself to have contributed scientifically to the article. According to SS (2015-06-24, in comments from co-authors, BGs Review) OS called him on the same day and explained in more detail the background to why he wanted to leave co-authorship. The reason was explained as due to a conflict between the research team in which OS worked and PM´s group. OS is also one of the complainants and in the formal notification to the Vice-Chancellor Prof Hamsten he stated that he first became aware of that the article contained several examples of scientific misconduct in connection with proofreading and that he then contacted the other authors and withdrew his co-authorship. OS contributions at the time when he was still co-author, according to "author contribution", was that he, along with RH and YZ, participated in the CT scanning. Of the four points that the notifiers state as scientific misconduct two of them involve CT scanning. Nowhere in the documents, can I see that OS at any point have raised grievances or objections to the CT examinations performed or presented. It is also noteworthy that OS has not taken part in or commented on the method or the results sections, incl Figure 8a, or on the interpretation of the CT scan image during the time the manuscript was completed. in SS comments (2015-06-24) he attach e-mail correspondence, which clearly shows that all authors, including OS, received the manuscript, figures and legends to figures with a call to "review, edit and approve it as necessary" before the manuscript was submitted for publication.

In the previous review (BG) OS was presented as the one that performed all the CT scan examinations, but he stands in "author contribution" together with RH and YZ as "involved" in the galley. This is somewhat surprising; why not better emphasize the individual profile in an article with so many co-authors? In SS comments (2015-06-24) he descibes that OS offered to perform the CT scan examinations of the grafted animals, that SS delivered an animal to him and was present througout the entire investigation and SS witnessed when OS inserted the probe into the esophagus. Furthermore SS states that OS got the opportunity to comment and correct the manuscript before it was submitted for publication. Honesty and reliability are two of the principles The Code prescribes and I think OS has deviated from

good practice by abandoning them in his communication with his co-authors. In performing CT scans, OS used animals and even if PM is assigned principal investigator on the animal ethics permit also the executor should have proper training in lab animal science and inform himself on the content of the current state of ethics. Whether OS had proper laboratory training at the time is unclear but he has deviated from good scientific practice by administering contrast agent via an oral probe since this procedure is not included in the current animal ethics permit.

Also YZ, JH and RR, together with SS, have been identified as having performed CT scan examinations. To which extent this has occurred, if they independently conducted or only attended is unclear.

Acknowledgements
Thanks to Magnus Hillman, PhD and researcher, Department of Clinical Sciences, Lund University for invaluable help with reviewing data handling, statistical analysis and presentation of generated data. Thanks also to Anders Forslid, associate professor and university veterinarian, Lund University for your contribution on knowledge about animal health status and ethics permit.

Abbreviations
ACTREM, Advanced center for translational regenerative medicine
ANOVA, analysis of variance
CD90, cluster differentiation 90
CLINTEC, Department clinical science, intervention and technology
CT, computed tomography (datortomografi)
DNA, deoxiribonukleinsyra
DICOM, digital imaging and communications in medicine
GFP, green fluorescent protein
IHC, immunohistochemistry
KI, Karolinska Institutet
s.c., subcutaneous
MSC, mesenchymal stromal cells
qPCR, quantitative real-time polymerase chain reaction
SD, standard deviation
SEM, standard error of the mean
SFS, Svensk författningssamling
SJVFS, statens jordbruksverks föreskrifter
VEGF, vascular endothelial growth factor

People:
BG, Bengt Gerdin
JH, Johannes Haag
MLL, Mei Ling Lim
OS, Oscar Simonsson
PJ, Philipp Jungebluth
PM, Paolo Macchiarini
RH, Rainer Heuchel
SS, Sebastian Sjöqvist
YZ, Ying Zhao