Mouse MApp: an application to monitor the welfare of mice

Overall aim

The aim of this Challenge is to develop an application (app) that uses artificial intelligence and machine learning to automatically detect changes in facial expression and/ or body condition, to improve the monitoring of mouse welfare.

Duration

Up to 18 months

Budget

Up to £100k

Sponsor(s)

AstraZeneca, GSK, CRUK Manchester Institute - University of Manchester, The Sainsbury Wellcome Centre (University College London) and Agenda Vets

Background and 3Rs benefits

Millions of mice are used worldwide in research each year (EU statistical report 2019). Ensuring any pain or suffering is kept to a minimum requires careful monitoring of the animals so that appropriate action can be taken, and humane endpoints implemented. Welfare is monitored using a panel of indicators such as the assessment of pain, body and coat condition, and the weight of the animal. Traditional methods of assessment based on monitoring changes in behaviour or clinical signs (e.g. weight loss) are time consuming and can have other limitations as they may not be specific to pain or a sensitive indicator of health under certain conditions. A number of scoring systems (e.g. the mouse grimace scale and body condition scoring) have been developed that provide simple, reliable and non-invasive measures for assessing welfare at the cage side (Langford *et al*., 2010; Leach *et al*., 2012; Ullman-Cullere and Foltz 1999), but there has not been widespread adoption.

The mouse grimace scale

All mammals communicate emotions through facial expressions and changes in these can provide a reliable and rapid means of assessing pain. The mouse grimace scale (MGS) is a graded measure for changes in facial expression related to pain (Langford *et al*., 2010; Leach *et al*., 2012) that has been adopted by some research groups and institutions to assess signs of pain following procedures. The

Further information including resources that support MGS use are available through the NC3Rs website.

MGS was developed based on a change in five facial actions units (FAUs) – orbital tightening, ear position, cheek bulge, nose bulge and whisker position – with each FAU scored on a three-point scale from not present (0), through to moderate (1) and severe (2). These action units increase in intensity as a response to post-procedural pain and can be used as part of a clinical assessment. Depending on the score, interventions can then be taken to alleviate pain and/or distress.

Grimace scores can be used to assess pain in real time at the cage side. Each animal has to be observed for a short period of time to avoid scoring brief changes in facial expression that are unrelated to pain and can only be carried out on awake animals. The MGS has been used for assessing both post-operative pain (Leach *et al*., 2012) and the effectiveness of analgesics (Matsumiya *et al*., 2012) and has the advantage that it does not require the mouse to be handled.

Automating the Mouse Grimace Scale

A number of projects have worked to automate the process of assessing pain using the MGS. In 2011, the Rodents Face Finder® software was published as a tool to automate the selection of images for scoring through detecting rodent eyes and ears in images, but the grimace scale scoring was still carried out manually (Sotocinal *et al*., 2011). Ernst *et al*., also successfully automated the process for pre-selecting images most suitable for manual scoring using an algorithm (Ernst *et* al., 2020a and 2020b). However, images selected by the algorithm are still evaluated manually. Tuttle *et al*., developed an automated MGS using machine learning and deep neural networks, which was shown to be highly accurate (94%) when compared to manual scoring (Tuttle *et al., 2018*). However, the current model only detects grimacing in albino mice and only provides a binary read-out (pain or no pain) rather than a precise MGS score. Andresen *et al*., developed an automated facial expression recognition software using deep learning neural networks to assess post-anaesthetic and/or post-surgical effects in mice (Andresen *et al*., 2020). Like Tuttle *et al*., the software only provides a binary read-out (e.g. post anaesthetic/surgical effect, no post anaesthetic/surgical effect) and has only been used in black-furred mice (C57BL/6JRj).

Body condition scoring

Body condition scoring (BCS) is a method to assess the welfare of mice without relying solely on measuring weight. Weight loss, measured as a percentage decline from initial weight or compared with the weight of age-matched controls, is commonly used as a criterion for welfare assessment. However, weight loss may not always be a sensitive indicator of animal health. For example, studies that create physiological changes, such as intraperitoneal fluid retention or tumour growth, may mask weight loss. BCS grades the amount of flesh covering bony protuberances on palpation or by visual assessment and correlates to potential changes in the health of the mouse (Ullman-Cullere and Foltz 1999). It uses clinical indicators which are scored as degree-of-deviation-from-normal, thereby allowing an animal to be monitored over time as health declines (Ullman-Cullere and Foltz 1999). Body condition is scored on a scale of one (emaciated) to five (obese). BCS is particularly useful for mice with tumours or ascites where changes in weight may be misleading. The BCS is typically carried out twice a day and requires the home cage to be taken off the rack and removal of the lid to

fully assess the animal including its movement and behaviour. The BCS can sometimes also require handling to palpate body condition, which can be stressful for the animal. There is currently no automated or semi-automated approach to BCS that would avoid the need to handle the mice.

Both the MGS and BCS are subjective, can be labour intensive and rely on the experience of the staff carrying out the observation, potentially leading to inconsistencies in scoring. They also require staff to be trained and how this training is given may vary. As a result, the MGS and BCS have not been widely adopted. There is a need to automate the process of scoring to deliver increased consistency and accuracy in welfare assessments, and to facilitate their widespread use.

This Challenge aims to create a facial and/or body recognition app based on artificial intelligence and machine learning that automates detection of pain and body condition in mice and is simple and fast to use in an animal facility setting. The worldwide availability of the tool could help improve and harmonise decisions regarding animal welfare and humane endpoints, reducing individual suffering in large numbers of mice used in scientific studies and procedures, while at the same time improving the the quality of scientific data when pain and discomfort can be mitigated. It could also reduce staff time and is likely to be adopted widely if it is quick, accurate and easy to use.

Key deliverables

The aim of this Challenge is to develop an app that uses artificial intelligence and machine learning to automatically detect changes in facial expression and/or body condition, to monitor the welfare of mice.

Real time images via a mobile device could be used to provide standardised read-outs of MGS and/or BCS scores. The solution must be cost effective, simple to use and easily accessible (e.g. a smartphone, tablet or laptop app) to enable wide-spread use and maximise the welfare benefits to mice.

Essential:

- An app or prototype app that can automatically recognise, read and score mouse facial expressions and body condition consistently and has the following features:
    - Works with mice of different coat and eye colours.
    - Is easy to use in an animal unit, is portable and requires limited training other than how to use the app (e.g. on a smartphone, tablet or laptop).
    - Provides secure storage of the images and data.
    - Is compatible, allowing image or data (i.e. score) download to other systems (e.g. computing infrastructure) to allow integration with other study data.

- Is reasonably priced to allow widespread uptake. For example, licenses can cost from £50 to £2k depending on whether the licence is per person or per organisation.

- Minimum read-outs are numerical indicators for either MGS or BCS that are comparable with current scales/ scores.

- Validate and compare to manual assessments to ensure the automated approach provides accurate MGS and BCS scores.

It is important to note that the CRACK IT Challenges competition is designed to support the development of new 3Rs technologies and approaches, which will improve business processes and/or lead to new marketable products. The application must include a plan to commercialise the results into a product or service. This should be taken into consideration when completing your application.

Sponsor in-kind contributions

The Sponsors will provide:

- Expertise in animal models and the use of MGS and/or BCS.

- Annotated images for MGS and BCS in order to support the development of automated assessments.

- The opportunity to test the developed product in both industry and academic settings.

References

2019 report on the statistics on the use of animals for scientific purposes in the Member States of the European Union in 2015-2017. [Online] [Accessed 23 July 2020].

Andresen N *et al*. (2020). Towards a fully automated surveillance of well-being status in laboratory mice using deep learning: Starting with facial expression analysis. *PLoS One* 15(4): e0228059.

Ernst L *et al*. (2020a). Improvement of the Mouse Grimace Scale set-up for implementing a semi-automated Mouse Grimace Scale scoring (Part 1). *Laboratory Animals* 54(1): 83-91.

Ernst L *et al*. (2020b) Semi-automated generation of pictures for the Mouse Grimace Scale: A multi-laboratory analysis (Part 2). *Laboratory Animals* 54(1): 92-98.

Langford DJ *et al*. (2010). Coding of facial expressions of pain in the laboratory mouse. *Nature Methods* 7(6): 447-449.

Leach MC *et al*. (2012) The assessment of post-vasectomy pain in mice using behaviour and the Mouse Grimace Scale. *PLoS One* 7(4): e35656.

Matsumiya LC *et al.* (2012). Using the Mouse Grimace Scale to reevaluate the efficacy of postoperative analgesics in laboratory mice. *Journal of American Association for Laboratory Animal Science* 51(1): 42-49.

Sotocinal SG *et al*. (2011). The Rat Grimace Scale: a partially automated method for quantifying pain in the laboratory rat via facial expressions. *Molecular Pain* 7: 55.

Tuttle AH *et al*. (2018). A deep neural network to assess spontaneous pain from mouse facial expressions. *Molecular Pain* 14: 1744806918763658.

Ullman-Cullere MH and Foltz CJ (1999). Body condition scoring: a rapid and accurate method for assessing health status in mice. *Laboratory Animal Science* 49(3): 319-323.