The ‘Real Welfare’ scheme: benchmarking welfare outcomes for commercially farmed pigs

F. Pandolfi†, K. Stoddart2, N. Wainwright2, I. Kyriazakis1 and S. A. Edwards1

1School of Agriculture Food and Rural Development, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK; 2AHDB PORK, Stoneleigh Park, Kenilworth, CV8 2TL, UK

Introduction

Several different groups in society take an interest in farm animal welfare with different perspectives taken (Fraser, 2003). Animal welfare is protected by legislation under which inspections are carried out annually (European Council, 1998). Additional safeguards are increasingly adopted through the mechanism of farm assurance schemes, which incorporate welfare standards and adopt third-party inspection procedures to verify compliance (Veissier et al., 2008). Historically, both legislation and assurance schemes have adopted resource-based measures of welfare but limitations appear when it comes to understanding the true welfare state of individual animals (Webster et al., 2004). For this reason, there has been a growing trend for the adoption of animal-based measures, sometimes called welfare outcome measures, which rely on measurements made directly on the animals themselves.

Implications

The ‘Real Welfare’ scheme aims to assess pig welfare on finishing farms using animal-based measures. This paper reports the results from the first 3 years of this scheme, based on the assessment of 5 463 348 pigs. This initiative represents the first long term, nation-wide benchmarking of welfare outcomes, and provides indications that such an approach can be practical for demonstrating good management and promoting welfare improvement. Its value rests on a demonstrable reduction of the prevalence of animal-based measures of welfare problems over the scheme’s application, and in the potential for similar schemes to influence on-farm welfare positively.
irrespective of their keeping conditions (European Food Safety Authority (EFSA), 2012). Such measures are now recognized as a better alternative to assess animal welfare across different environments (Whay et al., 2005). The application of this approach on farms was pioneered by the EU Welfare Quality® project (Blokhuis et al., 2010). Farmers also place great importance on animal welfare and perceive a relationship between good welfare and good animal performance (Hubbard et al., 2007). However, on-farm assessments of welfare outcomes are subject to many practical constraints, and must be quick, cheap and sufficiently flexible to adapt to different production systems and be meaningful for the end user (Edwards, 2007). Simplified versions, relying on so-called iceberg indicators, are consequently being investigated (Heath et al., 2014). Munsterhjelm et al. (2015a), by establishing the number and composition of possible sub-scales within the animal-based measures using principal component analysis, showed that different animal welfare issues could be captured with a short list of animal-based measures. The British pig industry has been very proactive in consideration of animal welfare and was the first to adopt Farm Assurance at a national level (Whittemore, 1995; Farm Animal Welfare Council, 2001). In 2006 they commissioned a project to investigate the feasibility of adopting welfare outcome assessments on British pig farms (Mullan et al., 2009a, 2009b, 2011a and 2011b). Following pilot studies, a protocol was adopted as part of the Red Tractor Assurance scheme for finisher herds. The objective of this paper is to report the prevalence of five main welfare outcomes for the mainstream finisher pig herds of the United Kingdom (excluding hospital pens) for the first 3 years of this scheme. This represents the first long term, nation-wide benchmarking of welfare outcomes for pigs – or any other species – on commercial farms at this scale. This study also describes the changes over calendar years of the different measures of welfare and the farm population involved through different variables related to farm environment and management.

### Material and methods

#### Data and data management

The data used in this analysis were collected from April 2013 to May 2016 in order to assess on-farm pig welfare through the ‘Real Welfare’ assessment scheme, as required for those finishing pigs under the Red Tractor Assurance scheme. The data were collected using a standardized protocol, owned and managed by the Agriculture and Horticulture Development Board. The welfare of the pigs was assessed by vets from 89 different veterinary practices carrying out quarterly health and welfare inspections for the Red Tractor scheme. The data are collected to inform the farm health plan. Although the welfare outcomes themselves are not audited by scheme providers, the completion of actions agreed between the veterinarian and the producer to address any issues is included in audits. Before undertaking the additional ‘Real Welfare’ audits, all vets underwent the same online and practical training in the assessment of the designated welfare outcomes (http://pork.ahdb.org.uk/health-welfare/welfare/real-welfare/real-welfare-vets/). The assessment involved five main measures (Table 1 and Table S1), chosen after stakeholder consultation to capture the most important welfare issues for the industry, using protocols developed and piloted in a previous research project (Mullan et al., 2009a, 2009b and 2011b).

#### Table 1 Measurements used in the assessment

<table>
<thead>
<tr>
<th>Measurements</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pigs requiring hospitalization</td>
<td>Pigs that would benefit from removal to a hospital pen</td>
</tr>
<tr>
<td>Yes</td>
<td>Pigs that would not benefit from removal to a hospital pen</td>
</tr>
<tr>
<td>Lame pigs</td>
<td>Pigs with signs of lameness</td>
</tr>
<tr>
<td>Lame</td>
<td>Pigs without any sign of lameness</td>
</tr>
<tr>
<td>Pigs with tail lesions</td>
<td>Pigs with severe tail lesions. Proportion of tail has been removed by biting</td>
</tr>
<tr>
<td>Severe</td>
<td>or tail is swollen or held oddly, or scab covering whole tip or fresh blood visible</td>
</tr>
<tr>
<td>Mild</td>
<td>Pigs with mild tail lesions</td>
</tr>
<tr>
<td>No lesions</td>
<td>Pigs without any of the above lesions</td>
</tr>
<tr>
<td>Dirty</td>
<td>Pigs dirty enough to obscure potential mild lesions</td>
</tr>
<tr>
<td>Pigs with body marks</td>
<td>Pigs with severe body marks extending into deeper layers of skin or lesions</td>
</tr>
<tr>
<td>Severe</td>
<td>covering a large percentage of skin</td>
</tr>
<tr>
<td>Mild</td>
<td>Pigs with mild body marks</td>
</tr>
<tr>
<td>No lesions</td>
<td>Pigs without any of the above body marks</td>
</tr>
<tr>
<td>Dirty</td>
<td>Pigs dirty enough to obscure potential mild body marks</td>
</tr>
<tr>
<td>Enrichment use</td>
<td>Pigs interacting with enrichment in the pen</td>
</tr>
<tr>
<td>Other</td>
<td>Pigs interacting with other pen features or pen mates</td>
</tr>
</tbody>
</table>

Each pig in the sample selected was classified into one of the several levels for each measurement (the classification for Enrichment use only concerns the active pigs of the sample). Detailed definition in Supplementary Table S1.
which assessed the sampling strategy, the interdependence, the variation and the reliability of the outcome measures. The measures were recorded from a sample of finishing pigs from the mainstream herd (i.e. excluding those in hospital pens). The number of pens assessed at each visit and the type of pens were selected to be representative of the farm. The sampling used was a multistage sampling. At the first level, all farms that belong to the Red Tractor Assurance scheme were sampled. At the second level, several pens were randomly selected on each farm in order to represent approximately one third of the pig places present in the farm. At the third level, all pigs in the pens were assessed for the prevalence of lameness and pigs requiring hospitalization. A random sample of pigs were further assessed for tail lesions and body marks (all pigs in the pen if there were fewer than 25 pigs, 25 pigs if there were up to 100 pigs, or 50 pigs if there were more than 100 pigs, and chosen to be representative of the pen) (Supplementary Material S1).

In addition to the welfare outcome measures, additional information about the sampled pens was also recorded during the visit: pen size (retrospectively categorized as small <30 pigs, medium ≥30 to <200 pigs, large ≥200); pen type (indoor kennels, indoor open pens with internal divisions, indoor open plan pens, Trobridge pens with an indoor and outdoor section, kennel + yard pens with indoor and outdoor areas, shelters in a field, other); ventilation type (natural, powered); feed form (liquid, meal, pellet); feed availability (ad libitum, restricted); feeder type (floor, hopper, trough). Enrichment was also recorded and retrospectively classified as being either a substrate or an object. From the five enrichment classifications, two substrates (straw, other substrate) and three object descriptors (chain (with or without attached object), plastics (e.g. hollow containers, tubes) and other object) could be selected. The quantity of straw could be assessed as restricted (portions dispensed throughout the day), low (<5 cm depth or <50% lying area covered), medium (depth of >5 cm over 75% of lying area) and deep (covers >75% pen floor, depth 30 cm+). Default qualification of the quantity of straw was used in case the quantity was not mentioned. Therefore, only the pens directly assessed by the vet without default classification were kept, leaving 74 596 pens with data on the quantity of straw. Only the farms with the mention ‘none seen,’ indicating the absence of visible enrichment in the pen at the time of the assessment, were considered as without enrichment. The mention ‘none seen,’ as distinct from a missing entry, was recorded only from June 2014 (sample of 76 002 pens).

The database was checked for mismatches and outliers. The different types of enrichment were transformed in dummy data in order to record the presence or the absence of each of the categories. From the date of the assessment, the calendar year and the season were extracted. Four seasons (Spring (March, April, May), Summer (June, July, August), Autumn (September, October, November) and Winter (December, January, February)) were identified from the date of assessment. All the measures reported in Table 1 were transformed into percentages, based on the total number of pigs assessed in the pen. Enrichment use was calculated as a ratio based on the following formula:

\[
\text{Enrichment use ratio} = \frac{\text{Number of active pigs interacting with the enrichment}}{\text{Number of active pigs interacting with pen features or pen mates or with the enrichment}}
\]

For the complete database, a sample of pens was assessed from 1928 farm units. In some cases one ‘farm unit’ could consist of farms at several different locations. Repeated measures were taken in the same farm units over time, giving 112 240 records at pen level. These concerned a total of 13 480 289 pigs present in the farm during the assessments, with 5463 348 pigs directly assessed using the ‘Real Welfare’ protocol.

Over the period of scheme implementation, the recording of tail lesions and body marks underwent some changes. After an initial 8-month period, a review of the functioning of the scheme decided that the recording of the enrichment use, minor tail lesions (dirty and mild tail lesions) and minor body marks (dirty and mild body marks) should become optional. However, the recording of the severe lesions continued to be mandatory. The vet could therefore decide to report either only the severe lesions or both the minor and the severe ones. The initial period from April 2013 to November 2013 included 9153 pen records from 1108 farms and the database over the 4 calendar years which included pens with recording of both severe and minor lesions and body marks included 28 247 pen records from 1293 farms.

Data analysis

Descriptive analysis of the farm characteristics and the welfare outcomes. Data processing and data analysis was carried out using Microsoft Access Office Professional Plus 2010, Microsoft Excel Office Professional Plus 2010 and RStudio for R-3.1.0 software for Windows (64 bit). The herd size of the farms was described at farm level. For all the farms a description was undertaken at pen level for the variables related to the environment, the feed and for the different types of enrichment, as these could vary within farm. In order to investigate the association of the type of enrichment and the different measures related to environment of the pigs, \(x^2\) tests or Fisher’s tests were used. A descriptive analysis was conducted for the percentage of pens and pigs with undocked tails and tails of different length. In order to better understand the confounding effect of tail docking and the different measures related to the environment of the pigs, \(x^2\) tests or Fisher’s tests were used. To assess the change of use of enrichment (substrates and objects) and the proportion of pens with undocked-tail pigs over years, generalized linear mixed models were used. In the first model, the binary variable was pens with undocked tails v. pens with mixed length tails or docked tails. The presence or absence of substrates in the pens was considered as the dependent variable.
variable in the second model and the presence or absence of objects was considered as the dependent variable in the third model. For these three models, the variable ‘year’ was considered as a fixed effect and the farm unit was considered a random effect. A descriptive analysis was conducted for the percentage of animals showing the different levels of each measure of welfare at farm and pen level. The pens in which the minor lesions were not recorded were excluded from calculations of the mean of the dirty and mild tail lesions and body marks. The variability between pens within the same farm was calculated as the intra-farm variance for the five main welfare outcomes (lame pigs, pigs requiring hospitalization, severe tail lesions, severe body marks, enrichment ratio use). The inter-pen and inter-farm variance was also calculated for the annual rolling average to provide a wider view of the differences, instead of focusing on one specific time point which might not reflect appropriately the welfare status in the farm.

Seasonal influences and annual averages of the welfare outcomes. The changes over calendar years of the different measures of welfare were assessed with a generalized linear mixed model in an analysis performed at pen level. The variable ‘year’ was considered as a fixed effect. The farm unit (Farm) was considered as a random effect as different pens could belong to the same farm. In order to reduce the information bias, the interaction between the veterinary practice that performed the assessment and the farm was also added as a random effect. Five different analyses were performed, considering the five main welfare outcomes as dependent variables. In order to identify the changes in the measures of welfare over the different seasons, the same analyses were performed for the variable ‘season.’ To look specifically at changes over time for farms initially having a higher prevalence of outcomes, farms with a prevalence of a specific welfare outcome above the 90th percentile in 2013 were selected separately according to each welfare outcome considered: lame pigs, pigs requiring hospitalization, pigs with severe tail lesions and pigs with severe body marks. As the values of the welfare outcomes were not normally distributed, a Friedman test was then used to assess the differences between years for these selected farms. Farm identification was used as a blocking variable. In order to understand whether individual farms showed consistency in welfare outcomes over years, Kendall’s tau-b correlations were calculated between the average percentages of each year for the main welfare outcomes.

Correlation between the measures of welfare. In order to understand the associations between the five main measures of welfare, the correlation coefficients between these measures were calculated. As data were not normally distributed, Spearman’s rank correlation coefficients were calculated for all the variables at pen level. The correlation of the minor lesions (mild and dirty) among themselves and with the five main measures of welfare was performed using the whole database, but excluding all the pens without any record of the minor lesions, and separately on the database of the start-up period (April to November 2013).

Results

Farm characteristics, enrichment and tail docking
The population of interest included mainly pigs raised indoors. The minimum herd size (pig places) was 12 and the maximum 24,000 with a mean of 1542. In all, 50% of the herds had 498 to 1586 pigs in the farm unit during the visit (1810 holdings ≥300 pig places). A breakdown of the housing and feeding practices in the study population is shown in Table 2. The three most common enrichment types were straw, chain and plastic (Table 3). Only 3.7% of the pigs had both enrichment types in the pens (substrates and objects) but this corresponds to 14.5% of the farms. Substrates were more common than objects with 62.0% of pigs (69% of the farms) with one or more substrates; and 31.9% of the pigs (52.5% of the farms)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Number of pens</th>
<th>Number of pigs assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pen type</td>
<td></td>
<td>DP</td>
</tr>
<tr>
<td>Indoors</td>
<td></td>
<td>DP</td>
</tr>
<tr>
<td>Kennels</td>
<td>11579</td>
<td>270676</td>
</tr>
<tr>
<td>Open + internal divisions</td>
<td>35252</td>
<td>1527574</td>
</tr>
<tr>
<td>Open plan</td>
<td>56767</td>
<td>3288664</td>
</tr>
<tr>
<td>Indoors and outdoors</td>
<td></td>
<td>DP</td>
</tr>
<tr>
<td>Trobridge</td>
<td>3984</td>
<td>84224</td>
</tr>
<tr>
<td>Kennel + yard</td>
<td>2088</td>
<td>66698</td>
</tr>
<tr>
<td>Outdoors</td>
<td></td>
<td>DP</td>
</tr>
<tr>
<td>Shelter + field</td>
<td>1942</td>
<td>198957</td>
</tr>
<tr>
<td>Other</td>
<td>585</td>
<td>26246</td>
</tr>
<tr>
<td>Missing values</td>
<td>443</td>
<td>309</td>
</tr>
<tr>
<td>Ventilation type</td>
<td></td>
<td>DP</td>
</tr>
<tr>
<td>Natural</td>
<td>83572</td>
<td>4570736</td>
</tr>
<tr>
<td>Powered</td>
<td>27385</td>
<td>830028</td>
</tr>
<tr>
<td>Missing values</td>
<td>1283</td>
<td>62584</td>
</tr>
<tr>
<td>Pen size</td>
<td></td>
<td>DP</td>
</tr>
<tr>
<td>Large (≥200)</td>
<td>6180</td>
<td>1863606</td>
</tr>
<tr>
<td>Medium (≥30 to 200)</td>
<td>65579</td>
<td>2406862</td>
</tr>
<tr>
<td>Small (&lt;30)</td>
<td>40481</td>
<td>1192880</td>
</tr>
<tr>
<td>Feed form</td>
<td></td>
<td>DP</td>
</tr>
<tr>
<td>Liquid</td>
<td>18161</td>
<td>521066</td>
</tr>
<tr>
<td>Meal</td>
<td>25649</td>
<td>853848</td>
</tr>
<tr>
<td>Pellet</td>
<td>68404</td>
<td>4088125</td>
</tr>
<tr>
<td>Missing values</td>
<td>26</td>
<td>309</td>
</tr>
<tr>
<td>Feed</td>
<td></td>
<td>DP</td>
</tr>
<tr>
<td>Ad libitum</td>
<td>101123</td>
<td>5211662</td>
</tr>
<tr>
<td>Restricted</td>
<td>11091</td>
<td>251377</td>
</tr>
<tr>
<td>Missing values</td>
<td>26</td>
<td>309</td>
</tr>
<tr>
<td>Feeder type</td>
<td></td>
<td>DP</td>
</tr>
<tr>
<td>Floor</td>
<td>1377</td>
<td>26161</td>
</tr>
<tr>
<td>Hopper</td>
<td>88910</td>
<td>4710744</td>
</tr>
<tr>
<td>Trough</td>
<td>21927</td>
<td>726134</td>
</tr>
<tr>
<td>Missing values</td>
<td>26</td>
<td>309</td>
</tr>
</tbody>
</table>
The percentage of pigs with undocked tails tended to be higher in pens with substrates (Table 4 and Table S7). Compared with 2013, a significant increase of pens with undocked-tail pigs was observed over time \( (P < 0.05) \) (Table 4). The data from 2016 only concern a part of the year and the changes for 2016 should be re-assessed after review of the data until the end of 2016.

### Descriptive analysis of the welfare outcomes

The descriptive analysis of the welfare outcomes (Table 5) shows some outcomes with high maximum values during individual visits of certain farms. However the median and upper quartiles both have very much lower values, highlighting the exceptional nature of these results. The descriptive analysis based on annual rolling averages also shows much smaller values (Table S8). The description at pen level of the welfare outcomes for the complete database and the start-up period is presented in the Supplementary Tables S9 and S10. The mean values of the intra-farm variance show much smaller values (Table S8). The description at pen level of the welfare outcomes for the complete database and the start-up period is presented in the Supplementary Tables S9 and S10. The mean values of the intra-farm variance shows much smaller values (Table S8).

### Trends over time

Compared with 2013, a significant decrease of the proportion of lame pigs and pigs requiring hospitalization was observed in 2014, 2015 and 2016 \( (P < 0.05) \). Compared with 2013, a significant increase of the proportion of pigs with severe tail lesions and severe body marks was observed in

---

**Table 3** Characteristics of the sample – number and percentage of pens and pigs with each enrichment type reported

<table>
<thead>
<tr>
<th>Enrichment</th>
<th>Percentage of pens with the enrichment of interest</th>
<th>Number of pens</th>
<th>Percentage of pigs assessed with the enrichment of interest</th>
<th>Number of pigs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Straw</td>
<td>44.7</td>
<td>50 136</td>
<td>60.8</td>
<td>3 320 398</td>
</tr>
<tr>
<td>Other substrates</td>
<td>1.41</td>
<td>1 588</td>
<td>2.46</td>
<td>134 313</td>
</tr>
<tr>
<td>Chain</td>
<td>24.2</td>
<td>27 196</td>
<td>16.4</td>
<td>894 112</td>
</tr>
<tr>
<td>Plastic objects</td>
<td>33.0</td>
<td>37 003</td>
<td>21.4</td>
<td>1 171 330</td>
</tr>
<tr>
<td>Other objects</td>
<td>8.92</td>
<td>10 014</td>
<td>7.09</td>
<td>387 608</td>
</tr>
<tr>
<td>Enrichment non seen</td>
<td>2.71</td>
<td>2058</td>
<td>1.73</td>
<td>65 613</td>
</tr>
</tbody>
</table>

1 Based on 76 002 pens and 3 790 879 pigs from June 2014 to May 2016.

**Table 4** Odds ratio, confidence intervals and \( P \)-values

<table>
<thead>
<tr>
<th>Year</th>
<th>Odds</th>
<th>95% CI</th>
<th>( P ) values</th>
<th>Odds</th>
<th>95% CI</th>
<th>( P ) values</th>
<th>Odds</th>
<th>95% CI</th>
<th>( P ) values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td></td>
<td></td>
<td>Intercept</td>
<td></td>
<td></td>
<td>Intercept</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>1.481</td>
<td>1.316</td>
<td>1.667 ( &lt; 0.001 )</td>
<td>1.811</td>
<td>1.723</td>
<td>1.902 ( &lt; 0.001 )</td>
<td>2.440</td>
<td>2.314</td>
<td>2.573 ( &lt; 0.001 )</td>
</tr>
<tr>
<td>2014</td>
<td>1.066</td>
<td>0.946</td>
<td>1.202 0.29</td>
<td>2.483</td>
<td>2.359</td>
<td>2.614 ( &lt; 0.001 )</td>
<td>2.139</td>
<td>2.027</td>
<td>2.257 ( &lt; 0.001 )</td>
</tr>
<tr>
<td>2015</td>
<td>1.318</td>
<td>1.120</td>
<td>1.551 ( &lt; 0.001 )</td>
<td>3.151</td>
<td>2.924</td>
<td>3.394 ( &lt; 0.001 )</td>
<td>2.749</td>
<td>2.546</td>
<td>2.968 ( &lt; 0.001 )</td>
</tr>
</tbody>
</table>

Absence of tail docking, and the presence of enrichment at pen level were the dependent variables and the year was the independent variable in a model that considered the effect of farm.
2014 but also in 2015 for the severe tail lesions \((P<0.05)\). However, no significant differences were observed in 2016 compared with 2013 for the proportion of severe tail lesions \((P>0.05)\) and a significant decrease was observed in 2015 and 2016 for severe body marks \((P<0.05)\). Compared with 2013, no increase of the enrichment use ratio was identified in 2014 \((P>0.05)\), but further increases were identified in 2015 and 2016 \((P<0.05)\) (Table 6). Any conclusion for 2016 needs to wait until the data for the full year are available.

Figures 1 and 2 show the monthly averages for the different welfare outcomes over the 36 months. The value of the 90th percentile was used to select the farms with the highest prevalence for each of the welfare outcomes in 2013 and the mean values for these selected farms in each subsequent year are reported in Supplementary Table S12 and Figure S1. The means for each welfare outcome for the group of farms selected decreased over years. The Friedman test showed significant improvement between years \((P<0.001)\) for all welfare outcomes. The Kendall’s tau-b correlation coefficient showed that some welfare outcomes were correlated by farm between two consecutive years \((\tau>0.3, P<0.05)\), but these correlations were weakened over longer periods, suggesting that farms changed their relative ranking over time, but that change could be slow for some parameters (Supplementary Table S13, S14, S15 and S16).

### Seasonal influence

Prevalence of lame pigs and pigs that would benefit from being in a hospital pen were significantly higher in Spring than in Summer, Autumn and Winter \((P<0.05)\). Prevalence of severe body marks was also significantly higher in Spring than in Autumn and Winter \((P<0.05)\) and a tendency \((P=0.09)\) for a lower prevalence of severe tail lesions was also observed in Summer. Compared with Spring, a significant increase in the enrichment use ratio was observed in Autumn and Winter \((P<0.05)\) (Supplementary Table S17 and S18).

### Correlation between the measures of welfare

The percentage of pigs that would benefit from being in a hospital pen was positively correlated to the percentage of lame pigs, and the absence of tail lesions was positively correlated with the absence of body marks \((P<0.05, R>0.3)\) (Supplementary Table S19). For the two periods considered (the start-up period from April to November 2013 and the total period from 2013 to 2016), the correlations of mild tail lesions and body marks were similar. The percentage of pigs with a dirty tail was positively correlated with the percentage of pigs with a dirty body (Supplementary Tables S20 and S21).
Discussion

Description of the population of interest and limitations

The objective of this study was to assess the welfare of pigs in commercial pig finishing enterprises in the United Kingdom (excluding hospital pens) through five animal-based measures and to assess the changes over time and season of these measures. The study also represented an upstream task to describe farm characteristics and welfare outcomes in preparation for future risk factor analysis. To our knowledge, the data collected represent the largest data set available on animal-based welfare measures for finisher pigs existing in the world. This scale necessitated use of many different vets for data collection, and Temple et al. (2013) reported the possibility of a lack of intra and inter-observer reliability in assessments repeated over the time. However, another study of Temple et al. (2012) showed that the inclusion of inter-observer effects did not impact on the outputs of the different measures, and the measures of lameness in pigs by trained observers showed consistency in the study of Main et al. (2000). The standardized procedure and the training provided to the individual vets was designed to minimize...
Quality® protocol (considering more than 10 lesions in two
2009; Temple et al., 2011 and 2012). In both the
countries where the Welfare Quality® animal welfare assess-
National Animal Disease Animal Service (NADIS) or from other
results obtained previously in the United Kingdom by the
purposes, the benchmarks can be compared with different
other countries of Europe (Whay et al., 2007; Courboulay et al.,
prevalence of
severe posture (1.1%) and gait problems (1.4%)) and to the
were slightly lower or comparable with the prevalences
the average prevalence of body marks at farm level was 0.26%
was 0.5% if both severe and mild lesions are considered, and
was 0.2%. The average prevalence of tail lesions at farm level
In this study, the average prevalence of lameness at farm level
was 0.2%. The average prevalence of tail lesions at farm level
was 0.5% if both severe and mild lesions are considered, and
the average prevalence of body marks at farm level was 0.26%
if only severe body marks were assessed. These prevalences
were slightly lower or comparable with the prevalences
reported by NADIS (2007–2011) (lameness (0.2 to 0.6%) and
severe and mild tail lesions (1.2%)), to the lameness prevalence
reported in United Kingdom by Kilbride et al. (2009) (mild to
severe posture (1.1%) and gait problems (1.4%)) and to the
prevalence of finisher pig lameness or tail biting reported in
other countries of Europe (Whay et al., 2007; Courboulay et al.,
2009; Temple et al., 2011 and 2012). In both the ‘Real Welfare’
and Welfare Quality® protocols, milder forms of lameness are
not recorded and pigs in hospital pens are excluded from study.
The prevalences reported therefore do not fully reflect the
overall welfare impact of lameness, but take account of the
way in which lame pigs are being managed on the farm.
A different definition of body wounds was used in the Welfare
Quality® protocol (considering more than 10 lesions in two
body zones or more than 15 in a single zone), but the definition
can be considered close to the definition of severe body marks
in this study (Supplementary Table S1). A lack of representa-
tiveness of the whole population of finisher pig farms in smaller
scale studies might explain the higher prevalence in some
reports, but it also raises the question of potential under-
reporting in large scale projects like those detailed above. This
highlights the importance of sustaining the motivation of
assessors in order to avoid under-reporting.

Changes over time
All welfare outcomes referring to lesions or sickness in the
mainstream herd (excluding hospital pens), except the tail
lesions, decreased over years. The reduction of the recorded
prevalence might be the result of a better management of
sick/injured pigs which have been moved to hospital pens.
Whether there is a real reduction of the prevalence, or
better management of hospital pens, it is known that bench-
marking of health and welfare measures can lead to greater
awareness and motivation to improve (Tremetsberger et al.,
2015). For the farms with initially higher prevalence of welfare
outcomes (above 90th percentile), the reduction for all
welfare outcomes also suggests improvement of the welfare
status following the implementation of the scheme. The
increase of use of some forms of enrichment over the years
showed some parallel trends with the decrease of the pre-
valence of welfare outcomes over the same period. Tail
lesions did not show significant reduction over time but the
complex interactions between enrichment provision and
prevalence of undocked tails will have influenced this result.
Enrichments might have been used post hoc to control
tail biting problems arising from other environmental or
management issues, particularly in undocked groups, so that
the substrate provision alone might not show a simple causal
relationship.

A number of the welfare outcome measures were
observed to show a significant seasonal difference, as was also identified in a Finnish study on animal-based
welfare-measures (Munsterhjelm et al., 2015b). This know-
ledge is important when designing sampling strategies for
farm audits. A decrease of the prevalence of physical injury in
autumn and winter, and over years, corresponded to an
increase of the interaction of the animal with the enrichment
during the same period. The association between these
changes needs to be more critically assessed in further study,
as it cannot be assumed that this relationship is causal
until proven.

Variability within and between farms
As mentioned in the studies of Temple et al. (2011 and 2013)
and Whay et al. (2007), animal-based measures of welfare
show variability both within and between farms. This high-
lights the importance of an appropriate sampling strategy.
The prevalence of welfare outcomes at farm level ranged
between 0 and 40.5%, but the extreme values were unusual
and the vast majority of the farms did not present any
problems or showed a very low prevalence. The reasons for the
variability seen intra- and inter-farm in animal-based mea-
sures of welfare need to be assessed through the identifica-
tion of risk factors that tend to increase the prevalence of
disorders in certain farms and understanding of the multi-
factorial impact of housing, nutrition and management
practices (Averós et al., 2010; Taylor et al. 2012).

Conclusion
The ‘Real Welfare’ initiative is a unique national industry
scheme designed to benchmark welfare outcomes on finishing
pig farms, promote welfare improvement and demonstrate
good management. The results from the first 3 years of the
scheme demonstrate a reduction of the prevalence of most
animal-based measures. Further research is needed to understand if this is attributable to better management of sick or injured pigs that have been moved to hospital pens or better attention to animal welfare. However, the baseline data provided highlight the value of this initiative, and the large database generated by the scheme will be a valuable source of information for future risk assessment investigations.

Acknowledgments
The authors thank the ‘Real Welfare’ project Steering Group and scheme participants for making these data available for analysis.

Supplementary material
To view supplementary material for this article, please visit https://doi.org/10.1017/S1751731117000246

References

‘Real Welfare’ scheme: pig welfare outcomes