

Accepted Manuscript

Inter-Observer Agreement And Sensitivity To Climatic Conditions In Sheltered Dogs' Welfare Evaluation Performed With Welfare Assessment Protocol (Shelter Quality Protocol)

Greta Veronica Berteselli, Laura Arena, Luca Candeloro, Paolo Dalla Villa, Fabrizio De Massis

PII: S1558-7878(17)30196-X

DOI: [10.1016/j.jveb.2018.09.003](https://doi.org/10.1016/j.jveb.2018.09.003)

Reference: JVEB 1182

To appear in: *Journal of Veterinary Behavior*

Received Date: 11 October 2017

Revised Date: 24 August 2018

Accepted Date: 28 September 2018

Please cite this article as: Berteselli, G.V., Arena, L., Candeloro, L., Villa, P.D., De Massis, F., Inter-Observer Agreement And Sensitivity To Climatic Conditions In Sheltered Dogs' Welfare Evaluation Performed With Welfare Assessment Protocol (Shelter Quality Protocol), *Journal of Veterinary Behavior* (2018), doi: <https://doi.org/10.1016/j.jveb.2018.09.003>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



1 **INTER-OBSERVER AGREEMENT AND SENSITIVITY TO CLIMATIC CONDITIONS IN**
2 **SHELTERED DOGS' WELFARE EVALUATION PERFORMED WITH WELFARE**
3 **ASSESSMENT PROTOCOL (SHELTER QUALITY PROTOCOL)**

4 Greta Veronica Berteselli ^a, Laura Arena ^{a,b}, Luca Candeloro ^a, Paolo Dalla Villa ^a, Fabrizio De
5 Massis ^a

6 ^a Istituto Zooprofilattico Sperimentale di Abruzzo e Molise "G. Caporale", Campo Boario, 64100,
7 Teramo Italy

8 ^b Facoltà di Medicina Veterinaria, Università di Teramo, SP18, 64100 Teramo, Italy

9 *Corresponding author: g.berteselli@izs.it

10

11
12 Greta Veronica Berteselli, DVM, Ph.D (corresponding author)
13 Istituto Zooprofilattico Sperimentale di Abruzzo e Molise “G. Caporale”,
14 Campo Boario, 64100, Teramo Italy
15 E-mail: g.berteselli@izs.it

16
17 Laura Arena, DVM
18 Istituto Zooprofilattico Sperimentale di Abruzzo e Molise “G. Caporale”,
19 Campo Boario, 64100, Teramo Italy
20 Facoltà di Medicina Veterinaria, Università di Teramo, SP18, 64100 Teramo, Italy
21 E-mail: larena@unite.it

22
23 Luca Candeloro, GDStat,
24 Istituto Zooprofilattico Sperimentale di Abruzzo e Molise “G. Caporale”,
25 Campo Boario, 64100, Teramo Italy
26 E-mail: l.candeloro@izs.it

27
28 Paolo Dalla Villa, DVM
29 Istituto Zooprofilattico Sperimentale di Abruzzo e Molise “G. Caporale”,
30 Campo Boario, 64100, Teramo Italy
31 E-mail: p.dallavilla@izs.it

32
33 Fabrizio De Massis, DVM
34 Istituto Zooprofilattico Sperimentale di Abruzzo e Molise “G. Caporale”,
35 Campo Boario, 64100, Teramo Italy
36 Dog, Welfare, Shelter, Protocol, Reliability, Sensitivity
37 E-mail: f.demassis@izs.it

38

Abstract

39 **Abstract**

40 The Shelter Quality Protocol (SQP) is a concise and easily implemented tool for assessing dog
41 welfare and to identify critical aspects of the shelter environment. A first version of the protocol has
42 been modified in order to improve its performance. The aim of this study was to evaluate the
43 reliability of the measures included in the second version of the protocol (SQP2) by testing the
44 inter-observer agreement between two independent assessors. We evaluated the sensitivity of
45 animal-based measures in detecting the shelter dogs' welfare outputs during two different seasons.
46 Ten Italian shelters were assessed contemporaneously by two assessors to determine the reliability
47 of SQP2 measures. Inter-observer agreement was evaluated using the Cohen's Kappa for qualitative
48 variables and Pearson' correlation for quantitative variables. The SQP2 was also applied twice
49 (January and August) by the same observer in five Italian shelters to evaluate the sensitivity of the
50 protocol to seasonal condition changes. The quantitative variables, "Number of animals
51 shivering/huddling" and "Number of animals panting" were analyzed by Wilcoxon test. Credible
52 intervals (95%) were calculated using a beta distribution for qualitative variables: "Body
53 condition", "Skin condition", "Dog cleanliness", "Signs of diarrhea", "Coughing", and "Lameness".
54 The level of agreement between the two observers on the qualitative variables such as body
55 condition, lameness, skin condition, was quite high, ranging from substantial (0.61-0.80) to almost
56 perfect (0.81-0.99). Inter-observer agreement was also significant with Pearson correlation
57 coefficients ranging from 0.51 to 0.92 (e.g.; curious = 0.74; sociable = 0.83; barking level = 0.61).
58 "Number of animals panting" and "Signs of diarrhea" showed a significantly difference between the
59 assessments ($p < 0.05$). Animals with lameness, coughing and inadequate body condition increased
60 in the winter season whereas the skin lesions increased during the summer, but not significantly.
61 The behaviors of shivering/huddling were observed too infrequently to be meaningfully analyzed.
62 Consistent inter-observer agreement exists in assessing dogs' welfare using the SQP2 confirming
63 the reliability of the measures included in the protocol. The SQP2 shows potential in detecting

64 changes in dogs' welfare outputs due to different climatic conditions. Further investigations are
65 required to confirm the sensitivity of selected measures to different seasons.

66 **Keywords:** dog; animal welfare assessment; shelter; protocol; reliability; sensitivity

67

68 **1. Introduction**

69 It is widely recognized that the shelter environment negatively affects animal welfare and poses a
70 challenge to most dogs. The main challenges include the unfamiliar housing systems, different daily
71 routines, changes in feeding regimen and type of food, unfamiliar sounds, smells and sights, social
72 deprivation, presence of several unfamiliar animals and humans, and the absence of an attachment
73 figure (Moesta et al., 2015; Tynes et al., 2015). Several studies have reported that the stress
74 associated with shelter environment can contribute to onset of behavioral problems, such as stress-
75 related aggression, abnormal or repetitive behaviors, anxiety and fear-related disorders (Tuber,
76 1999, Beerda et al., 2000; Hennessy et al., 2001; Hiby et al. 2006; Dalla Villa et al., 2013; Titulaer
77 et al., 2013). However, the behavioral responses of the dogs can vary depending on the stressors
78 (acute or chronic; physiological or psychological) and individual variability (e.g. genetics, age,
79 early life experiences, and the success or failure of previous responses to stress) (Horowitz, 2004;
80 Moesta et al., 2015). Moreover, the experience of dogs in shelters is of concern, not only in terms of
81 animal welfare, but also for its potential effect on the likelihood of adoption (Diesel et al. 2013;
82 Duffy et al, 2014). In the worst case, dogs remain in the shelters for the remainder of their life or, in
83 the countries where “no-kill” policy is not in force (e.g., USA, UK some regions of Spain), they are
84 euthanized if they are not adopted (Moesta et al., 2015). In Italy, however, euthanizing dogs is
85 forbidden by the law unless they are seriously ill, incurable or proven dangerous (Italian law
86 281/1991). The dogs can be hosted in long-stay facilities until the day of the adoption. Improving
87 the welfare of shelter dogs must be considered a primary goal of rehoming centers. One way to

88 improve welfare is to reduce the time dogs spend in the shelter environment. This could be achieved
89 through successful adoptions and a decreased return rate (Posage et al., 1998; Diesel et al., 2008;
90 Luescher and Medlock, 2009; Braun, 2011; Reid and Collins; 2015). The detrimental effects of a
91 shelter environment can be mitigated through providing adequate housing and management which
92 meet the dogs' ethological needs, and ensuring the highest quality of care in the case of long-term
93 sheltering (Miller and Zawistowski, 2015).

94 The Shelter Quality Protocol (SQP) was developed in response to the issues related to long-term
95 sheltering and it fills the existing gap in the assessment of shelter dog welfare. This gap is due, in
96 part, to the lack of and/or variation in regulatory frameworks defining minimum requirements for
97 shelters (Barnard et al., 2016). The SQP was designed to be concise and easy to implement in
98 assessing dog welfare (Barnard et al., 2016). This protocol was inspired by the Welfare Quality®
99 approach. In particular, it was built around the four principles of good feeding, good housing, good
100 health and appropriate behavior. Each one of these principles are composed of different welfare
101 criteria which in turn include different welfare measures (Welfare Quality®, 2009). Since welfare is
102 the outcome of multi-factorial effects, multiple variables need to be considered when applying the
103 protocol (Sherman, 2010; Barnard et al., 2016). For example, the criterion "Absence of prolonged
104 hunger" is composed of welfare measures: "Body condition" and "Feeding". The latter measure
105 includes "Type of diet", "Presence of special diets" and "Feeding regime". The measures were
106 selected to assess specific welfare criteria, reflecting management procedures (management-based
107 measures), housing environment (resource-based measures) and direct welfare outcomes (animal-
108 based measures). The animal-based measures are considered the best indicator of animal welfare
109 because they give a direct reflection on animal welfare state (EFSA, 2012; Kiddie and Collins,
110 2014). However, in some cases, resource- and management-based measures were maintained
111 because of the valuable information they provide to complement the animal-based measures
112 (Veisser et al., 2011). The SQP provides three different levels of assessment: i) measures taken at
113 shelter level, which encompass all management-based measures; ii) measures taken at pen level (

114 both resource- and animal-based measures) assessed by observing a random sample of pens and all
115 animals confined within; iii) measures taken at the individual level: all animal-based, assessed by
116 observing a sub-sample of dogs housed in the pre-selected sample of pens.

117 The approach of SQP can be considered innovative for companion animal welfare assessment.
118 Although the protocol was validated through field testing, which demonstrated its feasibility and
119 reliability, the SQP was modified in light of subsequent field application and feedback (Barnard et
120 al., 2016). The refined version of the Shelter Quality Protocol (SQP2) remained similar in structure
121 to the first version (SQP). The SQP2 was built following the Welfare Quality® principles (4) and
122 criteria (12) (Table 1).

123 The aim of this study was to assess the reliability of SQP2 measures. This was achieved by testing
124 the inter-observer agreement between two assessors who evaluated a sample of ten dog shelter
125 which were long-term confinement shelters. Reliability indicates the reproducibility of
126 measurements, in particular it is the degree to which a measure is free from errors and will therefore
127 yield the same results when repeated (Taylor and Mills, 2006; Martin and Bateson, 2007;
128 Thanasegaran, 2009). In addition, the SQP2 was tested in five shelters to evaluate the sensitivity of
129 selected animal-based measures in detecting the shelter dogs' outputs related to seasonal conditions.
130 Sensitivity concerns the ability of a tool (e.g., a protocol) or measurement to detect small but
131 important changes (Martin and Bateson, 2007).

132 **2. Material and Methods**

133 ***2.1 Shelter Quality protocol-second version (SQP2)***

134 In a previous study by Barnard et al. (2016), the SQP was proven to be a valid, reliable and practical
135 tool for assessing dog welfare in long-term shelters and for identifying critical aspects and welfare
136 risks in shelter management and environment. In order to improve the animal welfare measurements
137 and increase the applicability to persons with varying backgrounds (e.g., veterinarians and shelter

138 operators) (Taylor and Mills, 2006; Barnard et al., 2016) the SQP was reviewed and refined using
139 feedback obtained from its dissemination and field application. This resulted in the Shelter Quality
140 Protocol-second version (SQP2). The SQP2 contains 23 welfare measures, which mainly involve
141 animal-based measures as these allow for direct information on the welfare state of animals. The
142 welfare indicators within the protocol were selected to assess the above-mentioned principle and
143 criteria identified by Welfare Animal® Consortium (Blokhuys et al., 2010) (Table 1).

144 The refinements of the SQP included both the elimination and introduction of specific variables to
145 improve tool performance; changes in the level of assessment (i.e., from shelter to pen level or from
146 individual to pen level) to obtain more accurate information and reduce assessment time; and,
147 finally, the modification of the measurement approach to simplify the data collection (e.g.,
148 defensive and offensive aggression was merged in a single behavioral category). The refinements
149 are set out below in detail.

150 The variables “Nasal discharge” and “Dyspnea” included in the SQP were eliminated due to the low
151 prevalence (below 1 per cent) of expression resulting from previous on-field assessment of 29
152 shelters. “Morbidity” was also excluded due to the difficulty in collecting this information, shelter
153 managers not always are able to provide this information (Barnard et al., 2016).

154 The level of assessment of “Signs of diarrhea”, “Coughing” and “Evidence of pain” was altered
155 from the individual level to pen level in order to obtain a more accurate estimate. The assessor had
156 to record these variables by observing all the animals in the pen instead of taking a sample of
157 selected dogs.

158 In the SQP2 welfare measures “Barking level” and “Emotional state” were assessed at pen level
159 instead of the shelter level. The assessment of shelter dog emotional state was performed using a
160 Qualitative Behavioural Assessment (QBA). QBA is a method which focuses on the observation of
161 the whole animal and characterizes and quantifies the animal’s dynamic demeanor as an expressive
162 body language using descriptors such as “sociable”, “aggressive” or “anxious” et cetera
163 (Wemelsfelder, 2000). Arena et al. (2017), through a scientific approach (Free-Choice-Profiling

164 methodology) developed a list of terms specifically focused on shelter dogs emotional state. This
165 list was included in SQP2.

166 To obtain information about shelter turnover, the following variables were included: “Number of
167 dogs entered in shelter”, “Number of adopted dogs”, “Number of dogs returned after adoption”,
168 “Number of dogs returned to owner”.

169 In order to obtain information about the promotion of human-animal relationship in the shelter the
170 variables “Presence of qualified personnel for dog training”, “Presence of qualified personnel for
171 behavioral rehabilitation” were added. There is evidence that in shelters basic dog training can help
172 the development of mental activities and constitute the basis for an adequate human-animal bond,
173 which is important for successful adoptions. Similarly, dog behavioral rehabilitation facilitates the
174 human-animal relationship, improves dog welfare and increases the likelihood of adoption (Taylor
175 and Mills, 2007; Luescher and Medlock, 2009).

176 The variable “Shelter from adverse weather conditions” (sun, wind, rain) was introduced in the
177 SQP2 to replace the resourced-based measures “indoor/outdoor area”. This addition alleviated
178 issues which arose during the on-field assessment using SQP for the assessment of thermal comfort
179 (criterion “Good housing”) due to the different definitions of outdoor/indoor area established by
180 each Italian regional law.

181 The last refinement of SQP merged defensive and offensive aggression into a single category to
182 assess the reaction of dogs toward unfamiliar people in order to facilitate the interpretation of dogs’
183 behaviors (Barnard et al., 2016). These refinements are summarized in Table 2.

184 Before the on-field application, the two assessors were familiarized with SQP2 through teaching
185 materials (videos and photos).

186 ***2.2 Inter-observer agreement assessment***

187 An inter-observer agreement was performed to assess the reliability of the protocol measures after
188 refinement (SQP2). Ten Italian long-term shelters were evaluated by two different assessors

189 simultaneously and independently. The two assessors were female, aged between 30 and 40 years,
190 both veterinarians with specialization in applied ethology and animal welfare. Both assessors had
191 previous experience working with dogs and were made familiar with the SQP2 by training in field
192 and by video and photo support. Shelters were selected on the basis of the following inclusion
193 criteria: long-stay facility and manager availability to take part in the study. The shelters were
194 located in four regions of North and Central Italy: Trentino (1), Veneto (2), Emilia Romagna (6),
195 and Marche (1). The assessment was carried out according to the methodological procedures
196 described in the SQP2 (Shelter Quality protocol, 2017)

197 The dogs were assessed while housed in their home pen. The pens were selected using a shelter map
198 and on the basis of the number of dogs housed in each pen (taking sample size into account). The
199 sample size depended on the total number of dogs housed in the shelter at the time of visit (Table
200 3). The selected pens covered the different facilities. The sample of animals assessed at individual
201 level only included dogs over 6 months of age and those who had been housed in the shelter for 2
202 months or more. A maximum of three dogs per pen were assessed at individual level. The shelter
203 assessment was carried out in a single day.

204 As previously underlined, the measures were taken according to 3 scoring levels: shelter level (the
205 shelter was evaluated as a unit), pen level (each selected pen was evaluated as a unit and all animals
206 confined were observed irrespective of the total number of animals) and at the individual level
207 (each selected animal was evaluated as a unit). Except the management section (measures at shelter
208 level) that consisted of an interview with the shelter manager, the whole assessment was carried out
209 independently by the two assessors.

210 The assessment was carried out first at the pen level. The assessor stood in front of the pen, two
211 meters from the fence and without interacting with animals (unless this was required by the
212 protocol, e.g., short test to assess reaction towards human), recording the measures at pen level.
213 The individual animal-based measures were then recorded. Among these measures, a short
214 behavioural test was carried out to assess the dogs' reactions towards unfamiliar people. The test

215 was divided in two steps in order to record the dogs' reaction. First, the assessor approached the
216 fence, standing in front of the pen and ignored the dog for 30 seconds. Second, the assessor
217 crouched talking gently to the dog for 30 seconds. Finally, the assessor recorded the emotional
218 state of dogs in the pen by filling in the Emotional State Profile sheet.

219 After assessing the first pen, the assessor moved on to the next one, following the same procedure.
220 The assessment ended when the last pen was assessed (Shelter Quality, 2017).

221 Inter-observer agreement was evaluated using the Cohen's Kappa for qualitative variables; these
222 variables were all categorical. The Pearson' correlation was used for quantitative variables which
223 were discrete and continuous. Level of significance was set respectively at $\alpha \leq 0.0020$ and $\alpha \leq$
224 0.0025 , after applying the Bonferroni correction. For all analyses, z scores and p values were also
225 computed to indicate whether agreement was more than could be expected by chance alone
226 (Cohen, 1968). Table 4 summarizes the different variables with their score system.

227

228 *2.3 Seasonal sensitivity protocol assessment*

229 To better define the sensitivity of SQP2 to seasonal changes, five Italian long-term shelters were
230 assessed during winter and summer.

231 The SQP2 was applied twice (January and August) by the same assessor in five selected shelters
232 located in three regions of Central Italy: Marche (1), Abruzzi (3) and Molise (1). Shelter inclusion
233 criteria were the same as those used for the inter-observer agreement assessment. The assessments
234 were carried out following the same procedure described above. During the assessments,
235 temperature and humidity were recorded.

236 A random selection of dogs were used for each assessment because the shelter population and the
237 location of dogs in their pens could change in time. This sampling approach reduced possible bias
238 in the results by limiting the likelihood that the assessor could remember the scoring of the dogs
239 recorded during the first assessment (Barnard et al., 2016).

240 Since animal welfare may be affected differently between seasons, eight variables potentially
241 sensitive to changes in seasonal conditions were included in the analysis. These measures were
242 selected because animal-based measures can give direct information on dogs' output. They were
243 assessed according to the SQP2.

244 The variables selected were either quantitative - i.e., "Number of animals shivering/huddling",
245 "Number of animals panting" - or qualitative - i.e., "Body condition", "Skin condition", "Dog
246 cleanliness", "Signs of diarrhea", "Coughing" and "Lameness". The quantitative variables were
247 analyzed by Wilcoxon test. Credible intervals (CI) (95%) were calculated using a beta distribution
248 for qualitative variables.

249 The authors hypothesized that the likelihood of observing animals panting, with diarrhea, or with
250 skin lesions, would increase during summer season, whereas the likelihood of observing animals
251 shivering/huddling, coughing, with inadequate body condition, with lameness, and with dirty coat
252 would increase in the winter season. Statistical analyses were carried out using R V.2.15.3.

253 **3. Results**

254 ***3.1 Inter-observer agreement***

255 For the inter-observer agreement study, 222 pens and 710 dogs, living in the same pens, were
256 assessed over a population of 847 dogs hosted in 406 pens. A subsample of 365 dogs were selected.
257 After analyzing the qualitative variables, the Cohen's Kappa analysis showed a high level of
258 agreement between the two observers, ranging from substantial (0.61-0.80) to almost perfect (0.81-
259 0.99) for the majority of variables. As could be expected, perfect agreement was obtained for the
260 variable of "Type of drinkers" ($k=1$). The measures of "Shelter from rain" ($k=0.89$), "Age class"
261 ($k=0.89$), "Shelter from strong wind" ($k=0.88$), "Skin condition" ($k=0.84$) and "Lameness" ($k=$
262 0.82) obtained an almost perfect agreement. On the contrary, the variable "Active repetitive
263 behaviors" showed a fair agreement ($k=0.30$). It was not possible to calculate the correlation for the

264 variables of “Other compulsive behaviors”, “Air circulation”, “Shelter from excessive sun”,
265 “Evidence of pain”, “Coughing”, because of the lack of variability in the data (Table 5).
266 Analyzing the quantitative variables, the agreement among the two assessors was also significant
267 with Pearson correlation coefficients ranging from 0.51 to 0.92. In particular, the number of animals
268 subdivided into two categories (“N° of animals \leq 20kg”: 0.92; “N° of animals $>$ 20kg”: 0.91)
269 obtained high agreement (Table 5).
270 Level of agreement between the two assessors on the most QBA adjectives (9/13) was ranging from
271 substantial (0.61-0.80) to almost perfect (0.81-0.99) whereas the agreement of remaining adjectives
272 (4/12) was scored as fair (ranging from 0.41 to 0.60). The variables “N° of animals
273 shivering/huddling” achieved the perfect agreement. For the variable “N° of animals panting” the
274 correlation analysis could not be carried out due to the high homogeneity of the data. All results are
275 summarized in Table 5. All P values were significant ($P < 0.001$).

277 ***3.2 Seasonal sensitivity protocol assessment***

278 For climatic sensitivity protocol assessment, 244 pens and 612 dogs (304 in summer and 308 in
279 winter), living in the same pens, were assessed over a mean population of 935 dogs hosted in 398
280 pens. A subsample of 505 dogs were selected (254 in summer; 251 in winter). The recorded
281 temperature and humidity during the different assessments are summarized in Table 6.
282 In summer 20% (60/304) of dogs showed panting whereas no dogs showed this behavior in winter
283 (0/308). Number of animals panting significantly increased during the summer season compared to
284 winter season (Wilcoxon, $p = 0.0001$). In winter season only 1% (3/308) of dogs showed
285 shivering/huddling. No dogs showed this behavior in the summer (0/304). The behaviors of
286 shivering/huddling were observed too infrequently to be meaningfully analyzed.
287 The Beta distribution showed that the observation of “Signs of diarrhea” was statistically significant
288 in summer season (Summer: percentage of observations = 21.4%, CI: 0.15-0.29; Winter: percentage

289 of observations 7.6%, CI: 0.04-0.13). Although no statistically significant differences were
290 detected, “Body condition” (Summer: percentage of observations = 9.2%, CI: 0.06-0.13; Winter:
291 percentage of observations 12.4%, CI: 0.09-0.17) “Coughing” Summer: percentage of observations
292 = 0.8%, CI: 0.002-0.048; Winter: percentage of observations 3.2%, CI: 0.01-0.07), “Skin
293 condition” (Summer: percentage of observations = 8.3%, CI: 0.05-0.12; Winter: percentage of
294 observations 7.3%, CI: 0.04-0.11) and “Lameness” (Summer: percentage of observations = 2.4%,
295 CI: 0.01-0.05; Winter: percentage of observations 4.1%, CI: 0.02-0.07) seemed to support the
296 hypothesis of the authors showing results in the expected direction. The observations of animals
297 with lameness, coughing and inadequate body condition increased in the winter season whereas the
298 skin lesions increased during the summer. “Dog cleanliness” (Summer: percentage of observations
299 = 16.7%, CI: 0.12-0.22; Winter: percentage of observations 12.9%, CI: 0.09-0.17) didn’t show
300 statistically significance difference between the two assessments and the result was opposite to the
301 expected direction.

302 4. Discussion

303 SQP was considered an innovative approach to companion animal welfare assessment, particularly
304 for long-term shelter dogs. Its validity, reliability and feasibility were proven with a previous study
305 by Barnard et al. (2016). Since the feedback obtained from its application on field and its
306 dissemination, a refinement of the protocol aimed to improve its performance in assessing dogs’
307 welfare. The modifications made in SQP permitted the development of the second version of the
308 protocol (SQP2).
309 The consistent level of agreement obtained between two assessors evaluating a sample of ten
310 shelters highlighted that the changes made on SQP didn’t affect the tool performance and confirmed
311 the reliability of measures in the canine welfare assessment. Moreover, since the stressors may
312 differ between seasons, the climatic sensitivity of some animal-based measures suggested the
313 potential usefulness of the SQP2 in assessing changes in dogs’ welfare outputs.

314 On the basis of the results, some considerations of the SQP2 can be made. In particular, the variable
315 “Signs of diarrhea”, that in SQP2 is assessed at pen level instead of individual level, showed
316 substantial agreement and therefore allowed the assessment to be simplified. When group housed,
317 recording signs of diarrhea by considering individual animals does not allow to gain a proper
318 estimate because the presence of diarrhea cannot be associated to an individual dog. The presence
319 of liquid manure in group housing pens allows the detection of animals with potential clinical
320 problems (e.g., enteric disorders) and, consequently, the identification of which individual is
321 affected. This observation allows team member to carry out clinical examinations and treatment
322 (Sokolow et al., 2005; Newbury et al., 2010). This animal-based measure was also significantly
323 sensitive to climatic changes. In this study, the prevalence of diarrhea was found to be higher in the
324 summer; this could depend on the seasonality of gastrointestinal diseases such as intestinal parasites
325 (McCarthy and Moore, 2000; Fontanarrosa et al., 2006).

326 Although the other animal-based measures which were analyzed didn’t show statistically
327 significance in the seasonal comparison, results were consistent with the hypothesis of the authors.
328 Outbreaks of coughing and lameness may be increased by cold temperatures and high humidity. On
329 the other hand, the favorable environmental condition due to summer conditions can facilitate
330 ectoparasities’ presence which, in turn, could cause the outbreak of skin problems in animals
331 (Altizer et al., 2006). To better understand the seasonality of clinical conditions in shelter’s dogs,
332 further investigations are required.

333 Issues highlighted through the previous application and dissemination of SQP included the
334 challenges in assessing pen adequacy when both indoor and outdoor areas were present. The Italian
335 national framework law 281/1991 on companion animals and stray dog prevention does not provide
336 standards for the managing and keeping of dogs in shelters. Instead, this is defined at regional level.
337 Therefore, generating high variability in shelters nationally. Assessing such different housing
338 conditions was therefore challenging. This aspect was addressed in SQP2 by modifying the
339 assessment of the housing adequacy to encompass the ability of the shelters to house dogs from

340 adverse weather conditions. After this refinement, no difficulties were reported during the on-field
341 application of SQP2, as demonstrated by the high level of agreement.

342 Moreover, maintenance of body temperature is essential for positive animal welfare. Therefore,
343 each dog housed in the shelter must be provided with adequate thermal comfort at all times of year.
344 Seasonal variation in conditions, such as low or high temperatures, may increase stress (Miller and
345 Zawistoski, 2015). The SQP2 allows the assessment of changes in dog welfare due to these
346 different seasonal conditions. In particular, the measure “Number of animals panting” was proved
347 to be sensitive to seasonal changes, increasing during hot seasons. In contrast, the relevance of the
348 measure “Number of animals shivering/huddling” remains still unexplored due to the low
349 prevalence of animals showing these behaviors during the cold season. Further research should
350 include a higher number of individuals to be assessed in winter season to confirm its relevance in
351 canine welfare assessment.

352 Panting as well as shivering and huddling with pen-mates are physiological responses shown by
353 dogs in order to cope with extreme temperatures and can be considered an indication of thermal
354 discomfort. For this reason, the presence of animals showing these behaviors can highlight
355 inadequate housing and/or management which should be corrected accordingly (Rooney et al.,
356 2009). Although animal age, breed and overall health status can affect an animal’s tolerance of
357 temperature, generally the range from 20°C (68°F) to 30°C (86°F) of environmental temperature is
358 considered the “thermoneutral zone”. In this range the dog is able to maintain normal body
359 temperature without a change in metabolic rate (National Research Council, 2006). United States
360 Department of Agriculture (2013) suggested the indoor facilities’ temperatures must never fall
361 below 7.2°C (45°F) or rise above 29.5°C (85°F) for more than 4 consecutive hours.

362 The consistent agreement obtained in scoring QBA descriptors (anxious, relaxed, aggressive,
363 playful etc.) at pen level confirms that the fixed list of terms included in SQP2 allows a reliable
364 assessment of dogs’ emotional experience in a shelter environment (Walker et al., 2016; Arena et
365 al., 2017).

366

367 To have good welfare, shelter dogs, such as all domesticated animals, should experience more
368 positive (e.g., pleasure, happiness) than negative (e.g., fear, pain) emotions (Boissy et al., 2007).

369 The richness of the qualitative terms used gives to the assessment the power to address dynamic
370 aspects of welfare including subtle important differentiations, such as between “relaxation” and
371 “depression” or between positive and negative excitement (“excited” vs. “nervous”). From a whole-
372 animal welfare perspective, the aim is to capture larger patterns of expression and their context
373 through a large range of terms. However, it has been demonstrated that training significantly
374 improves inter-observer agreement levels ensuring both the interpretation of terms and the use of
375 the QBA scale (Minero et al., 2015).

376 “Barking level” obtained a moderate agreement. Excessive vocalization may be a sign of
377 frustration, distress or boredom (Rooney et al., 2009). The presence of subjects who vocalize
378 excessively might also have a detrimental impact on the other dogs housed in the pen (Petak, 2013).
379 Moreover, high noise levels in dog shelters may cause hearing damage and public disturbance
380 (Beelsey and Mills, 2010). For this reason, assessing barking level could indicate that acoustic
381 safety and noise mitigation strategies are required. Such strategies may include removing the
382 subject from the group, controlling visitors’ access in the pens’ area or building noise abatement
383 facilities (Coppola et al., 2006; Scheifele et al., 2012).

384 The only measure which showed a low level of agreement was “Active-repetitive behaviors”. In the
385 scientific community, there is controversy about the definition and the meaning of this behavioral
386 category. The terms “repetitive behaviors” and “stereotypies” are often used interchangeably.
387 Stereotypies are defined as repetitive and unvarying behaviors without apparent goal or function
388 (Mason and Latham, 2004). Some studies interpret repetitive behaviors in sheltered dogs as
389 indicators of compromised welfare, which may be related to stress and frustration or to confinement
390 environments (Hetts et al., 1992; Beerda et al., 1999; Beerda et al., 2000). Their presence is usually
391 observed in association with sub-optimal environmental conditions (Denham et al., 2014; Mason,

1991), although this also depends on individual variability (Denham et al., 2014; Overall, 2013). Mason and Latham (2004) underline that repetitive behaviors can be an individuals' strategy to cope with stress and can be correlated with good or neutral welfare. The current results from the measure "Active-repetitive behaviors" highlighted these problems with interpreting the occurrence of these behaviors. Improving the accuracy of the definitions used in the protocol could improve the reliability of this measurement. The reliability of the recording may be influenced by the clarity of definition of behavioral category or measurement (Kiddie and Collins, 2014). For each possible repetitive behavior shown by sheltered dogs it could be useful to specify, in addition to the definition of the behaviors, a threshold or the duration of repetition as indicated in other studies. For example, pacing: dog repeatedly (>3) paces around the pen in a fixed routine; wall bouncing: dog repeatedly (>3) jumps up pen wall from side to side; tail-chasing: dog chases tail (> 3) for reasons other than discomfort or grooming; chewing bars: dog repeatedly chews and bites at the wire of the pen (> 20 sec) (Hetts et al., 1992; Hubrecht et al., 1992; Beerda et al., 1999; Stephen and Ledger, 2005). The complex relationship between animal welfare and repetitive behaviors suggests that while these behaviors can be used as an indication of suboptimal welfare, they should never be used as the only measure of welfare (Mason and Latham, 2004).

5. Conclusion

The lack of uniformity in regulatory frameworks defining minimum requirements of shelters has hindered the development of a specific tool to assess dog welfare in long-term shelters. The SQP was created responding to this need and its validity, reliability and feasibility were proven (Barnard et al., 2016). This protocol has been the first tool which can be easily applied by people from different professions (veterinarians, competent authority, shelter manager, et cetera) and it identifies critical areas requiring intervention.

416 The feedback obtained from its dissemination and application necessitated the improvement of the
417 tool by developing a second version of the protocol (SQP2). The changes made in the protocol did
418 not modify its applicability or the reliability of the measures included. The good level of agreement
419 obtained in this study confirmed that SQP2 remains a useful tool for welfare assessment of dogs
420 housed in long-term shelters with broad areas of application (i.e., rank dogs' rescue and commercial
421 or breeding facilities according to the level of welfare they are providing). The SQP2 showed its
422 potential in detecting the impact of seasonal conditions on animal welfare. Particularly, some
423 measures included in the SQP2 such as "Presence of diarrhea" and "Number of animals panting"
424 showed their sensitivity in assessing changes in dog welfare due to different seasonal conditions.
425 Further investigations are required to confirm the utility of "Number of animals
426 shivering/huddling". Although "Lameness", "Coughing" and "Skin condition" showed results in the
427 expected direction but not significantly. These measures also require further research, for example,
428 including a larger sample size to explore more extensively their sensitivity to different seasons.

429

430 **Conflict of interest**

431 The authors declare that there are no known conflicts of interest associates with this publication and
432 the financial support for this project didn't influence its outcomes.

433 **Acknowledgments**

434 The authors wish to thank all shelters' staff involved for their participation and cooperation in this
435 research and for their enthusiasm. This study was a part of a project financially supported by The
436 Italian Minister of Health, Ricerca Corrente IZSAM 04/13 RC. A special thanks to Dr. Lauren
437 Smith and Dr. Helen Gray for their English revision. Finally, the authors are grateful to the
438 anonymous reviewers for his/her constructive comments and suggestions.

439 **Authors' contributions**

440 All authors conceived and designed the study. Berteselli and Arena, refined the protocol (SQP),
441 recruited the shelters, applied the SQP2 on-field in the selected shelters, collected the data and
442 drafted the manuscript. Candeloro performed the data analysis and advised on the study design, and
443 collaborated for the interpretation of the results. Dalla Villa was the responsible of operative unit for
444 animal welfare assessment and the supervisor of on-field activities. He also revised and approved
445 the final manuscript. De Massis was the supervisor on the entire project.

446 **References**

- 447 Altizer, S., Dobson, A., Hosseini, P., Hudson, P., Pascual, M., Rohani, P., 2006. Seasonality and the
448 dynamics of infectious diseases. *Ecol. Letters* 9(4), 467-484.
- 449 Arena, L., Wemelsfelder, F., Messori, S., Ferri, N., Barnard, S., 2017. Application of Free Choice
450 Profiling to assess the emotional state of dogs housed in shelter environments. *Appl. Anim. Behav.*
451 *Sci.* 195, 72-79
- 452 Barnard, S., Pedernera, C., Candeloro, L., Ferri, N., Velarde, A., Dalla Villa, P., 2016. Development
453 of a new welfare assessment protocol for practical application in long-term dog shelters. *Vet. Rec.*
454 178(1), 1-8.
- 455 Beaver, B. V., 2009. *Canine behavior: insights and answers*. Elsevier Health Sciences. 84.
- 456 Beerda, B., Schilder, M. B., Van Hooff, J. A., De Vries, H. W., Mol, J. A., 1999. Chronic stress in
457 dogs subjected to social and spatial restriction. I. Behavioral responses. *Physiol. Behav.* 66(2), 233-
458 242.
- 459 Beerda, B., Schilder, M. B., Van Hooff, J. A., De Vries, H. W., Mol, J. A., 2000. Behavioural and
460 hormonal indicators of enduring environmental stress in dogs. *Anim. Welf.* 9(1), 49-62.
- 461 Beesley, C. H., Mills, D. S., 2010. Effect of kennel door design on vocalization in dogs. *J. Vet.*
462 *Behav.: Clin. Appl. Res.* 5(1), 60-61.

- 463 Blokhuis, H. J., Veissier, I., Miele, M., Jones, R. B., 2010. The Welfare Quality project and beyond:
464 safeguarding farm animal well-being. *Acta Agric. Scand. A Anim. Sc.* 6, 129–140.
- 465 Bokkers, E. A. M., de Vries, M., Antonissen, I. C. M. A., de Boer, I. J. M., 2012. Inter-and intra-
466 observer reliability of experienced and inexperienced observers for the Qualitative Behaviour
467 Assessment in dairy cattle. *Anim. Welf.* 21(3), 307-18.
- 468 Boissy, A., Manteuffel, G., Jensen, M.B., Oppermann Moe, R., Spruijt, B., Keeling, L.J., Winckler,
469 C., Forkman, B., Dimitrov, I., Langbein, J., Bakken, M., Veissier, I., Aubert, A., 2007. Assessment
470 of positive emotions in animals to improve their welfare. *Physiol. Behav.* 92, 375–397.
471 <http://dx.doi.org/10.1016/j.physbeh.2007.02.003>.
- 472 Braun, G., 2011. Taking a shelter dog for walks as an important step in the resocialization process.
473 *J. Vet. Behav.: Clin. Appl. Res.* 6(1), 100.
- 474 Cohen, J., 1968. Weighted kappa: nominal scale agreement with provision for scaled disagreement
475 or partial credit. *Psychol. Bull.* 70, 213–220.
- 476 Coppola, C. L., Enns, R. M., Grandin, T., 2006. Noise in the animal shelter environment: building
477 design and the effects of daily noise exposure. *J. Appl. Anim. Welf. Sci.* 9(1), 1-7.
- 478 Denham, H. D., Bradshaw, J. W., Rooney, N. J., 2014. Repetitive behaviour in kennelled domestic
479 dog: stereotypical or not? *Physiol. Behav.* 128, 288-294.
- 480 Diesel, G., Pfeiffer, D. U., Brodbelt, D., 2008. Factors affecting the success of rehoming dogs in the
481 UK during 2005. *Prev. Vet. Med.* 84(3), 228-241.
- 482 Fontanarrosa, M. F., Vezzani, D., Basabe, J., Eiras, D. F., 2006. An epidemiological study of
483 gastrointestinal parasites of dogs from Southern Greater Buenos Aires (Argentina): age, gender,
484 breed, mixed infections, and seasonal and spatial patterns. *Vet. Parasitol.* 136(3), 283-295.

- 485 Hennessy, M. B., Voith, V. L., Mazzei, S. J., Buttram, J., Miller, D. D., Linden, F., 2001. Behavior
486 and cortisol levels of dogs in a public animal shelter, and an exploration of the ability of these
487 measures to predict problem behavior after adoption. *Appl. Anim. Behav. Sci.* 73(3), 217-233.
- 488 Hetts, S., Clark, J. D., Calpin, J. P., Arnold, C. E., Mateo, J. M., 1992. Influence of housing
489 conditions on beagle behaviour. *Appl. Anim. Behav. Sci.* 34(1-2), 137-155.
- 490 Hiby, E.F., Rooney, N.J., Bradshaw, J.W.S., 2006. Behavioural and physiological responses of dogs
491 entering re-homing kennels. *Physiol. Behav.* 89, 385-391.
- 492 Hubrecht, R. C., Serpell, J. A., Poole, T. B., 1992. Correlates of pen size and housing conditions on
493 the behaviour of kennelled dogs. *Appl. Anim. Behav. Sci.* 34(4), 365-383.
- 494 Italian law 281/'91, of 4 August 4th, 1991. Legge Quadro in materia d'affezione e prevenzione al
495 randagismo. *Gazzetta Ufficiale*, n. 203, 30 agosto 1991.
- 496 Kiddie, J. L., Collins, L. M. 2014. Development and validation of a quality of life assessment tool
497 for use in kennelled dogs (*Canis familiaris*). *Appl. Anim. Behav. Sci.* 158, 57-68.
- 498 Landis, J. R., Koch, G. G., 1977. The measurement of observer agreement for categorical data.
499 *Biometrics* 159-174.
- 500 Luescher, A. U., Medlock, R. T., 2009. The effects of training and environmental alterations on
501 adoption success of shelter dogs. *Appl. Anim. Behav. Sci.* 117(1), 63-68.
- 502 Martin, P., Bateson, P. P. G., Bateson, P., 2007. *Measuring behaviour: an introductory guide.*
503 Cambridge University Press.
- 504 McCarthy, J., Moore, T.A., 2000. Emerging helminth zoonoses. *Int. J. Parasitol.* 30, 1351-1360.

- 505 Miller, L., Zawistowski, S., 2015. Housing, husbandry, and behavior of dogs in animal shelters. In:
506 E. Weiss, H. Mohan-Gibbons, S. Zawistowsky (eds). *Animal Behavior for Shelter Veterinarians*
507 *and Staff*, 145-159
- 508 Minero, M., Dalla Costa, E., Dai, F., Murray, L. A. M., Canali, E., Wemelsfelder, F., 2015. Use of
509 Qualitative Behaviour Assessment as an indicator of welfare in donkeys. *Appl. Anim. Behav. Sci.*
510 174, 147-153.
- 511 Moberg, G. P., 2000. Biological response to stress: implications for animal welfare. In: G.M.
512 Moberg & J.A. Mench (eds), *The biology of animal stress: basic principles and implications for*
513 *animal welfare* CABI Publishing, Wallingford, 1-21.
- 514 Moesta, A., McCune, S., Deacon, L., Kruger, K. A., 2015. Canine enrichment. In: E. Weiss, H.
515 Mohan-Gibbons, S. Zawistowsky (eds), *Animal Behavior for Shelter Veterinarians and Staff*, 160-
516 171.
- 517 National Research Council, 2006. *Nutrient requirements of dogs and cats*. National Academies
518 Press.
- 519 Newbury, S., Blinn, M. K., Bushby, P. A., Cox, C. B., Dinnage, J. D., Griffin, BO'Quin, J., 2010.
520 *Guidelines for standards of care in animal shelters*. The Association of Shelter Veterinarians, 1-45.
- 521 Notari L., 2009. Stress in veterinary behavioural medicine. In Ed: Horwitz, D.F & Mills, D.
522 *BSAVA Manual of canine and feline behavioural medicine*. British Small Animal Veterinary
523 Association Publications, Gloucester. 136-145.
- 524 Overall, K.L., 2013. *Manual of Clinical Behavioral Medicine for Dogs and Cats*. Elsevier Health
525 Sciences, 231-309
- 526 Panel, E. A., 2012. Statement on the use of animal-based measures to assess the welfare of animals.
527 *EFSA Journal* 10(6), 2767.

- 528 Petak, I., 2013. Communication patterns within a group of shelter dogs and implications for their
529 welfare. *J. Appl. Anim. Welf. Sci.* 16(2), 118-139.
- 530 Phythian, C., Michalopoulou, E., Duncan, J., & Wemelsfelder, F., 2013. Inter-observer reliability of
531 qualitative behavioural assessments of sheep. *Appl. Anim. Behav. Sci.* 144(1-2), 73-79.
- 532 Posage, J.M., Bartlett, P.C., Thomas, D.K., 1998. Determining factors for successful adoption of
533 dogs from an animal shelter. *J. Am. Vet. Med. Assoc.* 213, 478-482.
- 534 Reid, P. J., Collins, K., 2015. Training and behavior modification for the shelter. In: E. Weiss, H.
535 Mohan-Gibbons, S. Zawistowsky (eds). *Animal Behavior for Shelter Veterinarians and Staff*, 172-
536 187.
- 537 Rooney, N., Gaines, S., Hiby, E., 2009. A practitioner's guide to working dog welfare. *J. Vet.*
538 *Behav.: Clin. Appl. Res.* 4(3), 127-134.
- 539 Scheifele, P., Martin, D., Clark, J. G., Kemper, D., Wells, J., 2012. Effect of kennel noise on
540 hearing in dogs. *Am. J. Vet. Res.* 73(4), 482-489.
- 541 Shelter Quality Protocol, 2017.
542 http://www.izs.it/IZS/Engine/RAServeFile.php/f/pdf_publicazioni/ProtocolloShelterQuality_EN_r
543 [ev24102017.pdf](http://www.izs.it/IZS/Engine/RAServeFile.php/f/pdf_publicazioni/ProtocolloShelterQuality_EN_r)
- 544 Sherman, B. L., 2010. Pets: behaviour and welfare assessment. In: Breed Md & Moore J (eds.)
545 *Encyclopedia of Animal Behaviour*. Oxford: Academic Press. pp 691-698
- 546 Sokolow, S. H., Rand, C., Marks, S. L., Drazenovich, N. L., Kather, E. J., Foley, J. E., 2005.
547 Epidemiologic evaluation of diarrhea in dogs in an animal shelter. *Am. J. Vet. Res.* 66(6), 1018-
548 1024.

- 549 Stephen, J. M., Ledger, R. A., 2005. An audit of behavioral indicators of poor welfare in kenneled
550 dogs in the United Kingdom. *J. Appl. Anim. Welf. Sci.* 8(2), 79-95.
- 551 Taylor, K. D., Mills, D. S., 2006. The development and assessment of temperament tests for adult
552 companion dogs. *J. Vet. Behav.: Clin. Appl. Res.* 1(3), 94-108.
- 553 Taylor, K. D., Mills, D. S., 2007. The effect of the kennel environment on canine welfare: a critical
554 review of experimental studies. *Anim. Welf.* 16(4), 435-447.
- 555 Thanasegaran, G., 2009. Reliability and Validity Issues in Research. *Integration & Dissemination*,
556 4.
- 557 Titulaer, M., Blackwell, E. J., Mendl, M., Casey, R. A., 2013. Cross sectional study comparing
558 behavioural, cognitive and physiological indicators of welfare between short and long term
559 kennelled domestic dogs. *Appl. Anim. Behav. Sci.* 147(1-2), 149-158.
- 560 Tuber, D. S., Miller, D. D., Caris, K. A., Halter, R., Linden, F., Hennessy, M. B., 1999. Dogs in
561 animal shelters: problems, suggestions, and needed expertise. *Psychological Science*, 10(5), 379-
562 386.
- 563 Tynes, V. V., Sinn, L., Koch C. S., 2015. The relationship between physiology and behavior in dogs
564 and cats. In: E. Weiss, H. Mohan-Gibbons, S. Zawistowsky (eds), *Animal Behavior for Shelter*
565 *Veterinarians and Staff*, 63-101
- 566 United States Department of Agriculture, 2013. *Animal Welfare Act*. Animal and Plant Health
567 Inspection Service, Title 7, Chapter 54.
- 568 Veissier, I., Jensen, K. K., Botrea, R., Sandøe, P., 2011. Highlighting ethical decisions underlying
569 the scoring of animal welfare in the Welfare Quality® scheme. *Anim. Welf.* 20(1), 89-.

- 570 Walker, J. K., Dale, A. R., D'Eath, R. B., Wemelsfelder, F., 2016. Qualitative Behaviour
571 Assessment of dogs in the shelter and home environment and relationship with quantitative
572 behaviour assessment and physiological responses. *Appl. Anim. Behav. Sci.* 184, 97-108.
- 573 Walker, J., Dale, A., Waran, N., Clarke, N., Farnworth, M., Wemelsfelder, F., 2010. The
574 assessment of emotional expression in dogs using a Free Choice Profiling methodology. *Anim.*
575 *Welf.* 19(1), 75-84.
- 576 Welfare Quality® 2009. Welfare Quality assessment protocol for pigs (saw and piglets, growing
577 and finishing pigs). Welfare Quality® Consortium: Leystad, NL
- 578 Wemelsfelder, F., Hunter, T.E.A., Mendl, M.T., Lawrence, A.B., 2000. The spontaneous qualitative
579 assessment of behavioural expressions in pigs: first explorations of a novel methodology for
580 integrative animal welfare measurement. *Appl. Anim. Behav. Sci.* 67, 193—215.
- 581

582 **Tables**

583 **Table 1: SQP2 measures associated with welfare principles and criteria.** Type of measures
 584 defined in brackets; management based measure (MBM); resources based measure (RBM); animal
 585 based measures (ABM). Measures were assessed according to different three units of assessment:
 586 the shelter (evaluate the shelter as a unit and all the animals within); the pen (evaluate the pen as a
 587 unit, taking in to account all of the dogs housed in the pen); the individual (evaluate each animal as
 588 a unit).

Principle	Welfare criteria	Welfare measure (type)	Sub-measures	Unit of assessment
Good feeding	Absence of prolonged hunger	Body condition (ABM)		Individual
		Feeding (MBM)	Feeding regimen Type of diet Special diets	Shelter
	Absence of prolonged thirst	Water supply (RBM)	Type of drinkers Availability of water Cleanliness of water Safety of drinkers	Pen
Good housing	Comfort around resting	Bedding (RBM)	Type of bedding At least one bed/dog Safety of bedding Cleanliness of bedding	Pen
		Safety of pen (RBM)	Sharp edges	Pen
		Cleanliness of animals (ABM)		Individual
	Thermal comfort	Thermoregulation (ABM)	Signs of thermal discomfort	Individual
		Shelter from adverse weather condition (RBM)	Shelter from excessive sun; wind; rain Air circulation	Pen
Ease of movement	Space allowance (RBM)		Pen	
Good health	Absence of injuries	Skin condition (ABM)	Presence of wounds; hair loss areas, swelling areas; evidence of parasites	Individual
		Lameness (ABM)		Individual
	Absence of disease	Evidence of pain (ABM) Signs of diarrhea (ABM) Coughing (ABM)		Pen
		Mortality (MBM)	Euthanasia for clinical problems; behavioral problem; Deaths (other than euthanasia)	Shelter
	Absence of pain induced by management procedures	Surgeries and control pain (MBM)	Presence of operating procedures for post-surgical monitoring; Presence of hospital pens Presence of protocol of analgesia	Shelter

Appropriate behavior	Expression of social behaviors	Social housing (MBM)	Single housing pens Pair housing pens Group housing pens (≤ 5) Group housing pens (> 5)	Shelter
	Expression of other behaviors	Abnormal behavior (ABM)	Repetitive and compulsive behaviors	Pen
		Barking (ABM)		Pen
		Exercise (MBM)	Exercise in outdoor areas Walking at leash	Shelter
	Good human-animal relationship	Reaction to human (ABM)		Individual
		Training and rehabilitation (MBM)	Presence of training personnel for activities with dogs, and specialized personnel in behavioral rehabilitation	Shelter
	Positive emotional state	Emotional state (ABM)	QBA	Pen

589

590

591

592 **Table 2: Refinements of SQP performed to develop the SQP2**

Variable	SQP	SQP2
Nasal discharge	At individual level	Eliminated
Dyspnea	At individual level	Eliminated
Morbidity	At shelter level	Eliminated
Signs of diarrhea	At individual level	At pen level
Coughing	At individual level	At pen level
Evidence of pain	At individual level	At pen level
Barking level	At shelter level	At pen level
Emotional state	At shelter level	At pen level. The adjectives' list was also refined
Number of dogs entered in shelter	-	Added
Number of adopted dogs	-	Added
Number of dogs returned to owner	-	Added
Number of dogs returned after adoption	-	Added
Presence of qualified personnel for dog training	-	Added
Presence of qualified personnel for behavioral rehabilitation	-	Added
Thermal comfort (good housing)	Indoor/outdoor area	Shelter from adverse weather conditions" (sun, wind, rain)
Defensive and offensive aggression	Considered as two separate categories	Considered as a unique category

593

594

595 **Table 3: Sample size.** The sampling includes only dogs over 6 months age and that have been in

596 shelter for 2 months.

Total number of housed dogs	Number of dogs to assess
Up to 29	All dogs
30-59	30
60-89	40
90-139	50
140 over	60

597

598

599 **Table 4: Scoring system of quantitative and qualitative variables.**

Quantitative variables	Type	Score	Qualitative variables	Type	Score
Aggressive	CV	VAS from 0 to 125 mm	Active-repetitive behaviours	CatV	Y-N
Alert	CV	VAS from 0 to 125 mm	Adequacy of pen area	CatV	Y-N
Anxious	CV	VAS from 0 to 125 mm	Age classes	CatV	Y-N
Barking level	CV	VAS from 0 to 125 mm	Air circulation	CatV	Y-N
Comfortable	CV	VAS from 0 to 125 mm	Body Condition	CatV	Adequate Too thin Too fat
Curious	CV	VAS from 0 to 125 mm	Clean water	CatV	Y-N
Depressed	CV	VAS from 0 to 125 mm	Cleanliness	CatV	Y-N
	CV		Coughing	CatV	Y-N
Excited	CV	VAS from 0 to 125 mm	Dry/clean bedding	CatV	Y-N
			Evidence of pain	CatV	Y-N
Fearful	CV	VAS from 0 to 125 mm	Fear/aggression test	CatV	Sociable Only fear Offensive/ defensive aggression
Hesitant	CV	VAS from 0 to 125 mm	Lameness	CatV	Y-N
Playful	CV	VAS from 0 to 125 mm	One bedding/dog	CatV	Y-N
Relaxed	CV	VAS from 0 to 125 mm	Safe bedding	CatV	Y-N
Nervous	CV	VAS from 0 to 125 mm	Sharp edges	CatV	Y-N
Sociable	CV	VAS from 0 to 125 mm	Shelter from adverse weather conditions (sun, wind, rain)	CatV	Y-N
Number of animals > 20Kg	DV	Ordinal number	Signs of diarrhoea	CatV	Y-N
Number of animals ≤ 20Kg	DV	Ordinal number	Skin condition	CatV	Y-N
Number of animals shivering/huddling	DV	Ordinal number	Type of bedding	CatV	Y-N
Number of animal painting	DV	Ordinal number	Type of drinkers	CatV	Y-N

600 CV: continue variable; DV: discrete variable; CatV: categorical variable; VAS: visual analogue scale; Y-N:yes-no

601

602 **Table 5: Inter-observer agreement**

Quantitative variables	Pearson' correlation Cor.P	Qualitative variables	Cohen's Kappa K
Aggressive	0,72*	Active-repetitive behaviours	0,30*
Alert	0,60*	Adequacy of pen area	0,85*
Anxious	0,60*	Age classes	0,89*
Barking level	0,61*	Body Condition	0,83*
Comfortable	0,74*	Clean water	0,66*
Curious	0,74*	Cleanliness	0,70*
Depressed	0,51*	Dry/clean bedding	0,60*
Excited	0,65*	Fear/aggression test	0,83*
Fearful	0,83*	Lameness	0,82*
Hesitant	0,60*	One bedding/dog	0,65*
Playful	0,70*	Safe bedding	0,64*
Relaxed	0,74*	Sharp edges	0,72*
Nervous	0,67*	Shelter from rain	0,89*
Sociable	0,84*	Shelter from strong wind	0,88*
Number of animals > 20Kg	0,91*	Signs of diarrhoea	0,77*
Number of animals ≤ 20Kg	0,92*	Skin condition	0,84*
Number of animals shivering/huddling	1*	Type of bedding	0,77*
		Type of drinkers	1*

603 *z-score, $P < 0.001$. Level of agreement in according to Landis and Kock (1977): 0.00, less than chance
604 agreement; 0.01-0.20 slight agreement; 0.21-0.40 fair agreement; 0.41-0.60 moderate agreement; 0.61-0.80
605 substantial agreement; 0.81-0.99 almost perfect agreement; 1 perfect agreement.

606 **Table 6: Temperature recorded during the assessments.**

Shelter region	Winter		Summer	
	Temperature	Humidity	Temperature	Humidity
Abruzzi*	9°C	74%	24°C	61%
Marche	9°C	76%	27°C	42%
Molise	10°C	70%	28°C	50%

607 *For Abruzzi the average of winter and summer temperatures and humidity were calculated over the three
608 shelters' assessments.

Highlights

- We refined the validation of dog welfare assessment protocol Shelter Quality (SQP)
- We assessed the Reliability of SQP measures through inter-observer agreement
- We assessed the Sensitivity of SQP animal-based measures to climatic changes
- Diarrhea and panting were sensible measures in assessing welfare between seasons
- Level of agreement on lameness, body condition, skin condition was high