

DAP 2020 HLES data: Document about Year Variables

2021-04-28

Contents

1	Introduction	1
2	Out-of-range year values	2
2.1	HLES Health Conditions	2
2.2	HLES Cancer Conditions	4
3	Diagnosis year before birth year	5
3.1	HLES Health Conditions	5
3.2	HLES Cancer Conditions	6

1 Introduction

This document provides information about two known issues regarding the DAP 2020 HLES Year variables:

1. Out-of-range year values that affect `hs_diagnosis_year` in the HLES Health Conditions data and `hs_initial_diagnosis_year` in the HLES Cancer Conditions data.
2. Instances where the diagnosis year of a health/cancer condition is before the birth year of the dog, which concerns `hs_diagnosis_year` in the HLES Health Conditions data, `hs_initial_diagnosis_year` in the HLES Cancer Conditions data, and `dd_age_years` in the HLES Dog_Owner data.

2 Out-of-range year values

We have identified instances where the respondent inputted values for the diagnosed year of a condition that are outside of the valid range of 1980-2025 allowed by REDCap (e.g., “07”). In these situations, REDCap simply presents a warning popup to respondents indicating they entered a value that is outside of the suggested range but still allows them to enter it and continue with the survey.

In addition, a minority of participants also specified a diagnosis year that is in the future (i.e., 2021-2025). Given that it doesn’t make sense to have a diagnosis year of a health/cancer condition that is in the future, the analysis in this report is based on limiting the year range from 1980 to 2020.

We have not modified these year values and are leaving it up to the researcher to either drop these out of range values or infer and correct them as they see fit.

2.1 HLES Health Conditions

In the 2020 HLES Health Conditions table, there are a total of 76884 records of reported health conditions that have a diagnosis year. Out of these, there are 275 instances where `hs_diagnosis_year` is out of the range of 1980-2020. The summary table below presents the out-of-range values and the number of times they appear in the data for respondents who provided a diagnosis year.

hs_diagnosis_year	count
0	5
1	15
2	12
3	2
4	1
5	3
6	10
7	2
8	1
9	2
10	3
11	2
12	3
13	3
14	3
15	6
16	3
17	6
18	18
19	29
20	38
21	1
29	1
108	1
111	2
112	1
153	1
199	1

hs_diagnosis_year	count
200	2
201	1
202	3
207	1
209	1
212	1
218	1
220	2
1015	1
1820	2
1916	1
1917	1
1918	1
1919	1
2021	10
2022	3
2025	2
2028	1
2029	1
2104	1
2105	1
2116	1
2819	1
2820	3
20014	1
20020	1
20083	1
20108	1
20111	1
20123	1
20183	1
20184	1
20188	2
20200	1
22016	1
22020	2
62016	1
82018	1
82020	1
102010	1
120111	1
151915	1
162015	1
182018	1
200505	1
200909	2
201414	1
201616	1
201919	2
202020	1
220200	1
1201111	1

hs_diagnosis_year	count
1202016	1
1920199	1
2014014	1
2020002	1
2020920	1
11121420	1
16201516	1
20102013	1
20111111	1
20162015	1
20201820	1
20202020	2
22017017	1
62006200	1
111111111	1
180180208	1
202002020	1
1620162015	1
2020202020	1
2200990008	1
2201501515	1
20111111111	1
20190192019	1
2017201702017	1

2.2 HLES Cancer Conditions

In the 2020 HLES Cancer Conditions table, there are a total of 1751 records of reported cancer conditions that have a diagnosis year. Out of these, there are 9 instances where **hs_initial_diagnosis_year** is out of the range of 1980-2020. The summary table below presents the out-of-range values and the number of times they appear in the data.

hs_initial_diagnosis_year	count
19	1
20	2
2021	1
2025	1
20202	1
90920	1
192019	1
2020200	1

3 Diagnosis year before birth year

We have identified instances in the 2020 HLES Health Conditions data and 2020 HLES Cancer Conditions data where the dog owner reported a diagnosis year for a health/cancer condition that is before the reported birth year of the dog. For example, for `dog_id` 90161, the reported diagnosed year for the Ear/Nose/Throat condition is 1980, but the birth year for the dog is 2013.

We have not modified these mismatched year values. We are simply documenting this issue and are leaving it up to the researcher to address these records as they see fit.

3.1 HLES Health Conditions

There are 494 records out of a total number of 76884 records in the 2020 HLES Health Conditions table where the dog owner reported a diagnosis year for a health condition that is before the reported birth year of the dog. Note that this is the total number of records after removing records with out-of-range values for the diagnosis year (i.e., records with a diagnosis year before 1980 and after 2020 were removed).

The table below lists the `dog_id`, `hs_condition`, `hs_diagnosis_year`, birth year (derived from either `dd_birth_year` in HLES_dog_owner or calculated by subtracting `dd_age_years` in HLES_dog_owner from 2020, the year the form was completed), and the difference between the diagnosis year and birth year for all of the 64 **skin condition records** where the diagnosis year comes before the birth year. Note that skin conditions start with the number 4 in the `hs_condition` column.

<code>dog_id</code>	<code>hs_condition</code>	<code>hs_diagnosis_year</code>	<code>birth_year_calculated</code>	difference
41831	498	1980	2008	-28
14015	498	1980	2008	-28
81633	404	1995	2009	-14
40360	423	2003	2015	-12
70684	406	1999	2009	-10
20001	430	2000	2010	-10
20001	432	2000	2010	-10
56738	432	2002	2011	-9
69567	411	2003	2012	-9
46652	409	2001	2009	-8
62396	426	2003	2010	-7
10440	426	2007	2014	-7
77863	411	1999	2005	-6
72786	406	2005	2011	-6
48516	407	2008	2014	-6
48516	426	2008	2014	-6
77863	410	2000	2005	-5
77863	412	2000	2005	-5
61575	411	2005	2010	-5
2591	406	2007	2012	-5
47802	412	2003	2007	-4
43844	405	2005	2009	-4
17405	498	2005	2009	-4
82125	426	2006	2010	-4
77127	423	2006	2010	-4
61575	432	2006	2010	-4

dog_id	hs_condition	hs_diagnosis_year	birth_year_calculated	difference
90060	408	2008	2012	-4
90060	412	2008	2012	-4
63687	432	2009	2013	-4
37543	414	2014	2018	-4
49450	414	2003	2006	-3
41068	430	2004	2007	-3
36130	412	2005	2008	-3
66910	408	2005	2008	-3
51744	423	2010	2013	-3
69389	412	2006	2008	-2
87299	411	2008	2010	-2
88383	428	2008	2010	-2
62510	406	2008	2010	-2
45984	498	2010	2012	-2
28684	423	2013	2015	-2
54591	404	2016	2018	-2
54591	498	2016	2018	-2
74760	429	2006	2007	-1
91905	423	2008	2009	-1
91905	426	2008	2009	-1
4886	426	2008	2009	-1
35718	412	2008	2009	-1
8514	415	2008	2009	-1
62491	412	2010	2011	-1
62491	423	2010	2011	-1
62491	426	2010	2011	-1
71133	426	2010	2011	-1
1644	426	2011	2012	-1
75450	406	2012	2013	-1
75417	406	2012	2013	-1
70198	411	2015	2016	-1
27713	410	2015	2016	-1
27713	425	2015	2016	-1
21587	426	2015	2016	-1
63300	430	2016	2017	-1
89501	404	2016	2017	-1
89501	412	2016	2017	-1
49207	498	2017	2018	-1

3.2 HLES Cancer Conditions

There are 8 records out of a total number of 1751 records in the 2020 HLES Cancer Conditions table where the dog owner reported a diagnosis year that is before the reported birth year of the dog. Note that this is the total number of records after removing records with out-of-range values for the diagnosis year (i.e., records with a diagnosis year before 1980 and after 2020 were removed).

The table below lists the `dog_id`, `hs_initial_diagnosis_year`, birth year (derived from either `dd_birth_year` in `HLES_dog_owner` or calculated by subtracting `dd_age_years` in `HLES_dog_owner` from 2020), and the difference between the diagnosis year and birth year for the records in the HLES Cancer Conditions table where the diagnosis year comes before the birth year.

dog_id	hs_initial_diagnosis_year	birth_year_calculated	difference
90814	2000	2008	-8
37605	2001	2009	-8
54936	2006	2012	-6
36563	2004	2008	-4
42124	2012	2015	-3
85582	2009	2011	-2
71059	2011	2013	-2
1281	2010	2011	-1