

**Input: Exclusions 2007-19 and 2020-22**

Records for previously recorded pupils 2007-19	208,915
Records for previously recorded pupils 2020-22	3,309
Records for new pupils 2020-22	16,685
<b>Total records 2007-22</b>	<b>228,909</b>

Previously recorded pupils 2007-19	79,372
Previously recorded pupils 2020-22	1,802
New pupils 2020-22	11,163
<b>Total pupils 2007-22</b>	<b>90,535</b>

Demographic keys for previously recorded pupils 2007-19	206,591
Demographic keys for previously recorded pupils 2020-22	3,309
Demographic keys for new pupils 2020-22	16,685
<b>Total demographic keys 2007-22</b>	<b>226,585</b>

Storage keys for previously recorded pupils 2007-19	207,782
Storage keys for previously recorded pupils 2020-22	3,309
Storage keys for new pupils 2020-22	16,685
<b>Total storage keys 2007-22</b>	<b>227,776</b>

**Output: Exclusions 2007-19\***

		<u>%Match Rate</u>
Records that matched to the spine in Exclusions dataset	195,862	93.8%
Individuals matched to the spine in Exclusions dataset	74,628	94.0%

**Output: Exclusions 2020-22\***

Records matched to the spine in Exclusions dataset	15,812	94.8%
Individuals matched to the spine in Exclusions dataset	10,581	94.8%

**Output: Exclusions 2007-22\***

Records matched to the spine in Exclusions dataset	214,876	93.9%
Individuals matched to the spine in Exclusions dataset	85,209	94.1%

\* These figures correspond to Unique Exact Matches (i.e., the stringent category of matches that is provided by default in research projects)

Dataset Name: Exclusions

Dataset ID: 2024-A003

### Degree of specificity for EAS datasets *(ADR-S Project Request)*

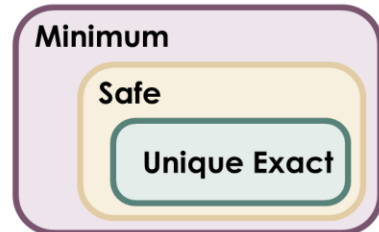
For Education datasets that are ADR-S ingested, **Exact / Safe / Minimum** are categories referring to the quality of the matches:

**Unique Exact** is the most stringent category, but with the highest precision. It only includes unique exact matches for sex, postcode, and date of birth (DEFAULT for projects). It excludes same sex twins.

**Safe** includes also high confidence matches: exact match for sex, exact match for date of birth, and partial match for postcode, this slightly increases the match rate, but also increases the risk of some incorrect matches

**Minimum\*** matches are those which meet the minimal threshold: exact match for sex, exact match for postcode, and two out of three matches from the three fields of date of birth (dd/mm/yy) - this offers the highest match rate, but introduces many false positives.

Competing matches (e.g., same sex twins) are included only as minimum. NRS strongly advises against their inclusion because it introduces many duplicate matches, involving incorrect links, that in many cases will be impossible for the researcher to resolve.



If a match is **Exact**, it'll be included as **Safe** and **Minimum** too; if a match is **Safe**, it will also be considered **Minimum**

\* *Minimum* matches have been called *Optimal* in the past. The name was changed as the word *Optimal* was misleading to represent the lowest quality of matches

The following figures correspond to a breakdown of the Unique Exact Matches (i.e., the stringent category of matches that is provided by default in research projects).

### **Exclusions A003 Summary tables**

#### **Sex**

		<b>Total%</b>
<b>Male</b>	7,482	71%
<b>Female</b>	3,099	29%
<b>Total:</b>	<b>10,581</b>	