Bradford District School Attendance Reporting Issues

Sam Relins, Dr Chris Brown, Prof Mark Mon-Williams

08/03/2022

Problems w. Data

As part of a study to evaluate air quality interventions in selected primary schools in Bradford, Bradford Institute for Health Research (BIHR) have requested aggregate school attendance data from the Bradford Local Authority. The data was subjected to simple quality assurance (QA) checks to validate the observations in the data, before beginning the full analysis and modeling exercise. This QA uncovered a number of issues with the data, that call into question the reliability of the current reporting process for school attendance.

This report gives a quick primer on school attendance data and then outlines the problems identified in the Bradford Local Authority School Attendance reports.

School Attendance Reporting

The following is a short primer on reporting standards for school attendance data - those familiar may wish to skip to the following section.

UK schools have a legal responsibility to collect attendance records for each registered pupil. The format and content of the attendance records is laid out in Department for Education (DfE) guidance [1]. The guidance divides the school week into 14 sessions - an AM/PM session for each day of the week, including weekends. Schools then record pupils' attendance in each session in according to a set of 24 standard "codes", defined by DfE and, which describe the attendance or absence of a pupil in educational activities. The following are a few examples of these codes:

- "/" and "": Present in school / = am = pm
- "L": Late arrival before the register has closed
- "V": Educational visit or trip
- "I": Illness (not medical or dental appointments)
- "O": Absent from school without authorisation
- "#": Planned whole or partial school closure (weekends for the Bradford schools in this study are recorded under this code, as none have planned weekend sessions)

Since the outbreak of the COVID-19 pandemic, DfE released further guidance [2] on attendance reporting, introducing a further 11 attendance sub-codes. These addictional sub-codes aimed to clarify attendance issues relating to the pandemic, for example:

- I01: Illness ("This code is used for pupils who are absent because of non-coronavirus (COVID-19) related illness or sickness")
- I02: Illness Confirmed case of coronavirus (COVID-19)
- X06: Pupil who is clinically extremely vulnerable if shielding is advised
- X07: Pupil advised specifically not to attend school as part of restrictions to education set out in Government advice

The various reporting systems employed by individual Bradford schools record this attendance data as an individual character, that represents the attendance code for that session e.g "/", "I", "L" and so on. The local authority then aggregate these attendance characters into a "string" or sequence of 14 characters

(e.g. "/\/\/\/\/\####"), on for each pupil on the register. These strings are then analysed or "parsed" to produce the aggregate attendance reports for the Bradford school district.

Problems with Local Authority Reporting

The QA of the attendance data began with a simple "sanity check": ensuring the total of the sessions reported under each attendance code was the same as the number of sessions that *should* have been recorded i.e. the total pupils on register multiplied by 14 weekly sessions. This ensures that all the required data is accounted for: that there aren't missing records for some pupils on the register or additional/duplicated records for others.

This check identified a number of issues with the attendance reporting from Bradford Council:

Duplicated entries:

Initial checks uncovered a large number of entries with a greater number of sessions than should have been possible. This was highlighted to the Bradford Council Information Management Team (IMT), and an investigation uncovered a system error that was duplicating student records, resulting in an overcount of certain attendance figures. These duplicate records were removed by the IMT and a revised report of the aggregate attendance figures was then produced.

Missing Reports:

After correcting for the duplicated records, a large number of the reported figures fell short of the total number of sessions that should have been reported. Assistance from the Bradford Council IMT uncovered a number of attendance codes that were omitted from the standard reports, which accounted for some of the missing data. There remained, however, a large number of records that had unaccounted or missing sessions.

Further investigation by the IMT identified issues with the process of aggregating the data from the raw attendance "strings" extracted from the individual information management systems at each school. 14 character attendance strings detail the attendance codes for each student for a week of attendance. The IMT identified a number of these strings that contained characters that couldn't be parsed or understood in their aggregation process - either the strings contained characters that weren't in the recognised DfE attendance codes, or the strings were too short. This resulted in missing session data for the affected entries in the attendance reports, amounting to hundreds of missing sessions at the aggregate level.

The IMT were unable to identify the source of these issues, and it was decided that individual schools would be contacted with the hope of comparing school-level records with those collected by the local authority.

Contradictions with raw data

A number of schools attempted a comparison of their own data with the data aggregated by the local authority, hoping to identify the cause of the missing attendance records. None were able to find an obvious explanation for the missing data. Raw data for the dates with the most missing records was then extracted and manually aggregated for comparison with the local authority reports. This analysis did not identify any missing records in the raw data extracts.

However, analysis of school-level attendance data uncovered yet more issues with the aggregate reports produced by the local authority. There was a lack of agreement on the basic calculation of the number of students on the register between the school's records and the local authority reporting. Several of the following aggregate attendance figures under the different DfE codes did not agree thereafter. The following is an extract from this analysis to illustrate the issue:

Source	observation	value
LA	total_pupils	175
Burley & Woodhead	total_pupils	205
LA	am	615
Burley & Woodhead	am	987
LA	pm	611
Burley & Woodhead	pm	978
LA	C	5
Burley & Woodhead	C	9
LA	I	9
Burley & Woodhead	I	0
LA	I01	10
Burley & Woodhead	I01	29
LA	X02	10
Burley & Woodhead	X02	14

These findings suggest that the aggregate reports of attendance produced by the local authority are missrepresenting the underlying data and should not be relied upon. The process of collecting and aggregating the attendance data should be reviewed if accurate reports are to be produced moving forwards.

Issues with Clarity of DfE Guidance

Aside from the problems with reporting at the local authority level, there were a number of issues with recording absences under the new COVID-19 guidance. Following the pandemic, the DfE introduced new guidelines dividing the existing "I - Illness" and "X - Not required to be in school" attendance codes into "sub-codes" that describe various COVID-specific attendance/absence scenarios. These guidelines were sufficiently ambiguous to result in a broad range of differing approaches when using these new sub-codes.

The principal source of confusion related to the new "Illness" sub-codes and, in particular, their designation designation as being "sub" codes. The existing "I" code was described as:

- Code I: Illness (not medical or dental appointments)

and the new COVID sub-codes were described as:

- Code I01: Illness - This code is used for pupils who are absent because of non-coronavirus (COVID-19) related illness or sickness.
- Code I02: Illness Confirmed case of coronavirus (COVID-19) - This code is used for pupils who are absent because they have tested positive for coronavirus (COVID-19).

This "sub" naming suggested to many administrators that the new codes are to be used in compliment with the existing "I" code, which remained available after the new COVID codes were introduced. In actuality, the system only allows the recording of one attendance code per session - so a session can be recorded as either "I" *or* one of "I01"/"I02".

As a result, many schools used all three "I" codes to record illness-related absences. The ambiguity between code "I" and code "I01" is problematic, as it isn't clear how absences recorded under these two codes differ from one another. Indeed, when asked, schools had many different approaches to recording absences under these three codes, examples including:

- "I" recording non-covid related illnesses, "I01" recording suspected but not confirmed cases of covid" and "I02" recording confirmed covid cases"
- "I" recording any illness absence as having occurred in the AM session, "I01"/"I02" recording the COVID/non-COVID nature of the illness in the PM session.

It seems that the guidance should have made clear that either:

1. the new "I"/"X" sub-codes were, in-fact, attendance codes in their own right, intended to replace the previous "I"/"X", or
2. How the new sub-codes should be used in conjunction with the existing "I"/"X" codes

It is also clear that a number of schools have noted the lack of "suspected COVID" illness scenario. The additional "X" sub-codes include "Code X02: Pupil self-isolating with coronavirus (COVID-19) symptoms", but it is unclear why this isn't an illness-related attendance code and, as a result, is rarely used in the attendance data. Indeed, several schools have attempted to use the "I"/"I01"/"I02" ambiguity to record these absences alongside non-covid and confirmed-covid cases.

In general, the lack of clear guidance introduces a subjectivity and a breadth of different approaches to reporting attendance between schools in the Bradford district. This causes significant problems when attempting to interpret these data on an aggregate level - indeed, it is likely that any existing analyses based on these data are innacurately reporting the true proportions of covid/non-covid related absences.

References

1. Aug 2020. **School attendance - Guidance for maintained schools, academies, independent schools and local authorities**, UK Department for Education, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1039223/School_attendance_guidance_for_2021_to_2022_academic_year.pdf

2. Jan 2022. **Recording non-attendance related to coronavirus (COVID 19)**, UK Department for Education, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1045090/Recording_non-attendance_related_to_coronavirus___COVID_19_.pdf