

An examination of the association between bullying, violence and crime in the ALSPAC cohort



Secondary Data Project Analysis Plan

Evaluating institution: University of Bristol

Principal investigator(s): Teyhan A, Cornish RP, Brennan, I

Analysis Plan for YEF Secondary Data Analysis Projects

Project summary

Project title ¹	An examination of the association between bullying, violence and crime in the ALSPAC cohort
Research Team	University of Bristol and University of Hull
Principal investigator	Alison Teyhan
Analysis plan author(s)	Jasmine Rollings, Alison Teyhan, Rosie Cornish, Iain Brennan, Bushra Farooq
Overarching research question ²	What is the association between: Bullying involvement (victimisation or perpetration) and: (a) self-reported and (b) official sanction for violent behaviour.
Supporting research question(s) ³	<i>Is there evidence for a dose-response relationship between bullying and outcomes, or of a sensitive period during which bullying is more strongly associated with later violence than at other time points?</i> <i>Is there a greater risk of outcomes for physical bullying involvement?</i> <i>What is the association between bullying involvement and self-reported weapon carrying?</i>

¹ Please make sure the title matches what's in the header.

² In simple terms written for a non-expert, what's the main thing this research projects sets out to answer?

³ What are the supporting research questions that will be tested in support of addressing the primary research question. This should not exceed three on the cover sheet and more detail can be added below if there are further questions to be addressed.

Dataset(s) to be used	- Avon Longitudinal Study of Parents and Children (ALSPAC) - Avon & Somerset Police (A&SP) Data - National Pupil Database (NPD)
Population characteristics	Our sample will be all individuals in ALSPAC for whom we have permission to link to crime and education data.
Years data spans	ALSPAC - participants born in 1991/92. A&SP data – available from 2007 to 2021 (small number of pre-2007 records).
Geographic coverage	Avon and Somerset, UK
Primary outcome(s) investigated	(1) Police record for serious violence (2) Self-reported violence
Main method(s) to be used or tested	To model the relationship between our exposures and outcomes we will use logistic regression models. Models will be adjusted for age and other confounders.

Analysis plan history

Version	Date	Reason for revision
1.X [<i>latest</i>]		
1.1		
1.0 [<i>original</i>]		

Any changes to the design or methods need to be discussed with the YEF. Describe in the table above any agreed changes made to the design.

Table of contents⁴

An examination of the association between bullying, violence and crime in the ALSPAC cohort	1
1.	5
1.1.	5
1.2.	7
1.3.	7
1.4.	8
2.	10
2.1.	10
2.2.	12
2.3.	15
2.4.	15
2.5.	15
3.	17
3.1.	17
3.2.	21
3.3.	22
3.4.	23
4.	23
4.1.	23
4.2.	24
Research questions: approach and methods.....	24
5.	27
5.1.	27

5.2. 28

6. 28

1. About the project

1.1. Background to the project

Successful violence interventions depend on the identification of upstream risk factors, an understanding of their relative importance, and how they relate to each other. In our current work we are examining several risk factors: adverse childhood experiences; positive educational factors; and school absence and exclusion. The project outlined in this proposal will build upon this other work by focusing on a risk factor that is associated with these other factors, and with violent offending - bullying.

Bullying is a common, worldwide social problem, which tends to peak in early to mid-adolescence as children develop greater independence, form new peer groups, and shape their adult identities and behavioural patterns. In England, 36% of Year 10 pupils (aged 14-15 years) in the 2014 Longitudinal Study of Young People in England reported they had been bullied [1]. Bullying can take many forms: physical; behavioural; verbal; or relational (e.g. disrupting social relationships between victim and peers). It can also be classified as 'overt' (e.g. direct physical or verbal attacks) or 'covert/relational' (indirect, such as passing notes or ostracising someone). Irrespective of the form, it has three key characteristics: repetition over time, harm, and unequal power [2]. It is thought that girls and boys experience similar prevalence of bullying, but the type differs, tending to be more overt and physical in boys [3]. Certain social groups are at greater risk of bullying victimisation, including children with special educational needs and disabilities. There can be an overlap between bullying and being bullied – 'bully-victims' are those who bully others and who are bullied themselves [3].

While the bullying itself may be time-limited, the negative effects may be long-term and persist throughout adolescence and into adulthood. This is potentially true for both victims and perpetrators of bullying. For example, bullies, victims, and bully-victims are all at increased risk of depression in adolescence [4, 5]. Outcomes observed in adults include an increased risk of substance misuse in those who were perpetrators [6], and mental health problems including anxiety and depression in those who were victims [4, 7, 8]. Victimisation is also associated with poor educational attainment and school absenteeism and may result in children leaving education at the earliest opportunity [4, 9].

Some studies have examined the relationship between bullying and later involvement in violent crime. Two meta-analyses have found that both bullying perpetration and victimisation are associated with weapon carrying [10, 11] and increased risk of later violence [12]. More recent longitudinal studies have found bullying perpetration in childhood and early adolescence to be associated with violence in later adolescence [8, 13] and adulthood [14]. Only two of the bullying and violence studies we have identified are UK-based [8, 13].

Weapon-carrying is a behaviour characterised by combined motivations to defend oneself from potential violent harm and to be at an advantage when doing violent harm [15]. These motivations mirror patterns in bullying where people can be perpetrators, victims or a combination thereof. In meta-analyses of the international literature, the strongest association with weapon-carrying is in bully-victims, but victims and bullies are also at heightened risk compared to those not involved in bullying [10, 11]. These studies included in these reviews were primarily from a US context with no studies from the UK, meaning the relationship between bullying and weapon-carrying in the UK is poorly understood.

Weaknesses in the existing literature

This existing bullying and violence literature has limitations which we believe we can address through our study:

- (1) Most studies only measure bullying at one timepoint, or if they have multiple timepoints they reduce the exposure measure to 'at any timepoint'. We propose examining whether the timing or duration of exposure is important.
- (2) Few studies disaggregate bullying behaviour into different types. ALSPAC data allows us to look at physical bullying separately from other types of bullying.
- (3) Most studies of bullying and violence focus on perpetrators. ALSPAC data will allow us to add nuance by considering both perpetration and victimisation and their intersection.
- (4) Some previous studies have been on boys only and/or had small sample sizes, both limitations our dataset overcomes.
- (5) Many studies only have self-reported outcomes. ALSPAC has both self-reported and official offending data. This linkage to crime data is an important strength as attrition from cohort studies is socially patterned, and participants who have the most troubled lives, including involvement in crime, will be at increased risk of non-participation but included in official data.

The pathway from bullying behaviours in childhood and early adolescence to violent behaviour in young adulthood could be causal or could reflect common factors that predispose young people to both being a bully or a victim and to later violence, such as abuse, negative peer group influences, low connectedness to school, and poor academic achievement [18-19]. ALSPAC has a wealth of measures at individual, family, and neighbourhood levels that would allow us to consider the role of confounding in the relationship between bullying and violence including family socio-economic position, childhood adversities, behavioural difficulties (as measured by the SDQ), and personality disorder (measured by the Childhood Interview for Borderline Personality Disorders).

By leveraging this unique and valuable data set, we will provide some of the most robust evidence to date about the pathways and predictors from bullying to violence and weapon-carrying, guiding the allocation of violence prevention resources and creating a better understanding of how this pervasive behaviour affects young lives.

1.2. Research question(s)

Our primary research question is:

- (1) What is the association between bullying perpetration or victimisation and (i) self-reported violent behaviour and (ii) official sanction for serious violence.

Secondary research questions are:

- (2) Does the association between bullying involvement and violence differ depending on (i) the duration of exposure to bullying involvement (i.e. is there a cumulative risk) or (ii) when the bullying involvement is experienced? (i.e. is there a sensitive period)?
- (3) What is the association between bullying involvement and self-reported weapon carrying in adolescence?
- (4) Does involvement with physical bullying have a greater associated risk of outcomes than other types of bullying?

Table 1.2. How will the questions be addressed at each stage?

Question Number ⁵	Interim report	Final report
1-4	Our interim report will include an overview of the data set and initial descriptive results summarising the datasets overall and our derivation of the study samples. We will also present descriptive statistics for the exposure, outcome and potential confounder measures.	Our final report will finalise the descriptive results, and then present the final results for each of the research questions in turn. It will include a discussion of the policy implications of the results and provide new insight into the relationship between bullying and violence.

1.3. Hypotheses

Based on findings from previous studies, we expect that the children in our study sample who have been involved with bullying will be at increased risk of violent behaviour and weapon-carrying in late adolescence/early adulthood. However, any observed relationship between bullying and violence could be a result of confounding. Potential confounders that we will consider include deprivation (both family and neighbourhood); poor engagement with education; friendships; behavioural difficulties; personality disorder; birth month; and adverse childhood experiences.

1.4. Key concepts

Table 1.4 Definitions of key concepts

Terms	Definition used
Self-reported violence at 17 or 18 years	ALSPAC participants were asked questions relating to antisocial behaviour and crime (in the past 12 months) in questionnaires or study clinics at ages 17 and 18 years. Serious violence was defined as saying yes to: (i) Hit/kicked/punched someone else on purpose with the intention of really hurting them? or (ii) Carried a knife or other weapon for protection or in case it was needed in a fight? Self-reported violence will be coded into a binary variable (yes/no). Yes will be coded if participants have answered 'yes' to either violent behaviours at either or both timepoints. No will be coded if participants have answered 'no' to both violent behaviours at both timepoints.
Self-reported weapon carrying at 15 years	Using the same phraseology as above, ALSPAC participants were asked at approximately 15.5 years of age, whether they had 'Carried a knife or other weapon for protection or in case it was needed in a fight?'. A binary variable will be derived to indicate if a participant responded 'yes' or 'no' to this question.
Official sanction for serious violence age 13-24 years	Home Office offence codes will be used to identify police records for: (i) violence against the person, indictable only; (ii) robbery, indictable only; and (iii) possession of weapons, triable either way or indictable only. [As per Home Office definition of serious violence]. We will examine offences that occurred between ages 13 and 24 years.

Bullying	At ages 8.5, 10.5 and 12.5 ALSPAC participants answered a modified version of the Bullying and Friendship Interview Schedule (BFIS) (Wolke <i>et al</i>, 2000, 2001a, 2001b; Woods & Wolke, 2003). Participants answered if a bullying event (received or given) had occurred in past 6 months (details of what constitutes these 'events' is set out below). If a child responded 'Yes' to any bullying event which had happened to/by them, a series of follow-on questions were asked. This included the frequency with which each event took place (infrequently: 1-3 times in past 6 months; frequently: more than 4 times in last 6 months but less than once a week; very frequently: at least once a week).
Bullying victimisation	We will use previously used definitions: a participant will be classified as a bullying victim if they received any of the components of either type frequently (several times a month) or very frequently (several times a week). OVERT RECEIVED ● Had personal belongings taken ● Been threatened/blackmailed ● Been hit/beaten up ● Been tricked in a nasty way ● Been called bad/nasty names RELATIONAL RECEIVED ● Others wouldn't play with them to upset them ● Been made to do things didn't want to ● Had lies/told nasty things said about them ● Had games spoilt Four variables will be derived: at each of the three time points a binary variable will measure whether or not the participant was a bullying victim, then a fourth variable will indicate if a participant was victimised at any of the three timepoints..
Bullying perpetration	As above, we will used previous definitions: participants will be classified as a bully if they had done any of the following to another child frequently or very frequently:

	OVERT GIVEN • Taken personal belongings from others • Threatened/blackmailed others • Hit/beaten up others • Tricked others in a nasty way • Called others bad/nasty names RELATIONAL GIVEN • Wouldn't play with others to upset them • Got others to do things didn't want to • Told lies/ said nasty things about others • Spoilt other children's games Four variables will be derived: at each of the three time points a binary variable will measure whether or not the participant was a bully, then a fourth variable will indicate if a participant was a bully at any of the three timepoints.
Physical bullying victimisation/perpetration	From the variables above there is one question that pertains to physical bullying victimisation (been hit/beaten up) and perpetration (hit/beaten up someone). Two binary variables will be derived to indicate whether a participant was physically bullied (victimisation) or physically bullied others (perpetration) at any timepoint (only those reporting frequently or very frequently will be marked as yes, in order to ensure this measure reflects the repeated behaviour that defines bullying).

2. About the datasets

2.1. Overview of datasets used

This project will use data from the Avon Longitudinal Study of Parents and Children (ALSPAC), which has been linked to education data and to local police records.

Avon Longitudinal Study of Parents and Children (ALSPAC)

ALSPAC is a birth cohort study – it has followed the same group of people from before they were born (i.e. during their mother's pregnancy) through to the present day. ALSPAC recruited pregnant woman who had an expected due date between April 1991 and December 1992 and who lived in a defined area in and around the city of Bristol, UK. There were approximately 14,000 study children alive at one year of age. Participants have been followed

up regularly via questionnaires and clinics. When the study children reached 18 years they were sent 'fair processing' materials that described ALSPAC's intention to link to their health and administrative records and gave them a clear means to object. The study children (now adults in their early 30s) have been followed throughout their lives via questionnaires and clinic visits, and through record linkage.

National Pupil Database (NPD)

ALSPAC has been linked to the National Pupil Database, which is a central repository of education data for children attending school in England. Individual-level education data available includes attainment, special educational needs, and free school meal (FSM) eligibility.

Avon and Somerset Police (A&SP)

ALSPAC has also been linked to Avon and Somerset Police data, which includes records of charges, cautions and other out of court disposals for crimes committed in Avon and Somerset (a geographical area than includes the original ALSPAC recruitment area). Data are available from 2007-2021. A pilot linkage of ALSPAC to the PNC showed that 86% of the crimes committed by the ALSPAC participants were in the A&SP area. Of the 12,662 ALSPAC participants for whom we have permission to link to their crime records, 2,273 have at least one A&SP record. 933 of the records, belonging to 520 ALSPAC individuals, relate to serious violence as defined by the Home Office.

The Police National Computer (PNC)

As part of this project, we are applying to link ALSPAC to the PNC: if this linkage is successful within the time-frame of the project, we will re-run our analyses using outcomes measured in the PNC. The PNC is a large database that contains information on offenders in England and Wales (E&W). Linkage to the PNC would provide the following benefits:

- (1) **National data.** The PNC would allow us to include all offences in E&W, not just those in A&S.
- (2) **Conviction data.** The PNC contains information on convictions, the A&SP data does not.
- (3) **Offences pre-2007 and complete offence histories.** PNC records are not weeded and cover the lives of the ALSPAC participants. Local police weed their records, meaning some older records for minor offences will not be included in our dataset, and A&SP data do not include pre-2007 records as these were paper-based.

Note that as we do not currently have the PNC data, we have not included the PNC in Section 2.2 below.

2.2. Secondary data source(s)

Table 2.2a Dataset Description - *ALSPAC*

Name of dataset	ALSPAC
Data owner(s)	University of Bristol
Type of data	Longitudinal prospective birth cohort study
Availability of data	The ALSPAC Executive operate a <u>managed open access</u> process.
Team member(s) who will have access	Alison Teyhan, Rosie Cornish, Senior Research Associate (SRA)
Population/geographic coverage or sampling frame	ALSPAC recruited pregnant woman who had an expected due date between April 1991 and December 1992 and who lived in a defined area in and around the city of Bristol, UK. ALSPAC recruitment took place in the old administrative county of Avon, UK. The catchment area covered the three health administration districts within the South-West Regional Health Authority that became the 'Bristol & District Health Authority'. This area includes the City of Bristol and the surrounding urban and rural areas.
Years covered or survey waves	1991 onwards
Exclusion criteria	Those who have died or withdrawn from the study. Those whose data is not available in UKSeRP.
Expected population/sample size (following exclusion criteria)⁶	There are approximately 14,700 children alive at one year whose data is available in UKSeRP. After applying the further exclusion criteria for the NPD and A&SP data (see boxes below), we would expect a sample size of around 9-10,000. Complete case analyses

⁶ This may not be known at this stage of the project. Please provide your best estimate or range based on your knowledge of the dataset.

	will have a lower sample size due to missing questionnaire and clinic data.
Documentation	Cohort profile papers: <u>Child</u> <u>Mother</u> <u>Website</u> (includes link to searchable data dictionary)

Table 2.2b Dataset Description – *National Pupil Database*

Name of dataset	National Pupil Database
Data owner(s)	Provided by the Department for Children, Schools and Families (DCSF) (now known as Department for Education).
Type of data	Administrative data.
Availability of data	The NPD data that has been linked to ALSPAC is available through ALSPAC's <u>managed open access</u> process.
Team member(s) who will have access	Alison Teyhan, Rosie Cornish, SRA
Population/geographic coverage or sampling frame	The NPD is a central repository of education data for children attending state school in England.
Years covered or survey waves	This project will use pupil level census data from Key Stage 1 (KS1) and the start of Key Stage 2 (KS2, age 7). The ALSPAC sample span three academic years. Participants started KS2 in academic years 1998/99 to 2000/01.
Exclusion criteria	Those who have not been given an opportunity to opt out of linkage to administrative data and those who have explicitly dissented to linkage to education data. ALSPAC study children who were not at a state school in England during these years will not be in this dataset (e.g. those who are home educated, privately educated, attending a school outside England).

Expected population/sample size (following exclusion criteria)⁷	We expect around 10-12,000 ALSPAC participants will have education data.
Documentation	Documentation for the education data that has been linked to ALSPAC is included in the ALSPAC Data Dictionary, which can be accessed from the ALSPAC Website . More information can be found on the NPD website .

Table 2.2c Dataset Description – Avon and Somerset Police Data

Name of dataset	Avon and Somerset Police Data
Data owner(s)	Provided by Avon and Somerset Police, extract linked to ALSPAC is owned by University of Bristol.
Type of data	Police recorded crime (charges, cautions, and other out-of-court disposals).
Availability of data	The A&SP data that has been linked to ALSPAC is available through ALSPAC's managed open access process.
Team member(s) who will have access	Alison Teyhan, Rosie Cornish, SRA
Population/geographic coverage or sampling frame	The A&S police data includes records of charges, cautions and other out of court disposals for crimes committed in Avon and Somerset (a geographical area that includes the original ALSPAC recruitment area).
Years covered or survey waves	2007-2021 (very few pre-2007 records due to paper records being used at that time).
Exclusion criteria	Those who have not been given an opportunity to opt out of linkage to administrative data; those who have explicitly dissented to linkage to criminal justice data; those living outside the A&SP area between ages 13 and 24 years.

Expected population/sample size (following exclusion criteria) ⁸	9-10,000
Documentation	A <u>data note</u> describing the linkage of ALSPAC to A&SP data has been published.

2.3. Primary data collection

No primary data will be collected.

2.4. Linking datasets

Linkage of ALSPAC to the NPD and to the A&SP data has already been achieved. An application to link ALSPAC to the PNC is in progress.

2.5. Access and data protection

The data for this project has been provided by an ALSPAC data manager and includes ALSPAC clinic and questionnaire variables, education variables (from the NPD), and crime variables (from A&SP). The variables were prepared for analyses and combined into one dataset. Therefore, the rest of this section refers only to this one dataset and not the three datasets listed above.

1. Data Protection

ALSPAC adheres to the principles of the ONS 'Five Safes framework' (safe data, safe projects, safe people, safe settings, and safe outputs).

Safe Data

In common with all projects using ALSPAC data:

- (1) the dataset was minimised to only include variables necessary for this project;
- (2) the data was de-identified;
- (3) the ID variable in the dataset was unique to this project meaning it could not be linked to any other ALSPAC data;
- (4) potentially disclosive variables (e.g. small cell counts, precise dates) were not released to researchers.

⁸ This may not be known at this stage of the project. Please provide your best estimate or range based on your knowledge of the dataset.

Safe projects

A proposal for this project was submitted to the ALSPAC Executive in October 2023 for approval. It was approved In November 2023 and assigned 'B number' B4443, and is now listed on the study's Proposal Summaries [webpage](#).

Safe People

All members of the research team are researchers experienced in working with sensitive, individual-level data. The data for this project was only accessible by three members of the research team who are all ONS accredited safe researchers (AT, RC and JR). AT, RC and JR have a DBS certificate dated within the last 24 months.

Safe setting

Due to the sensitive nature of the linked education and crime data, the dataset was accessed via [UKSeRP](#), a secure and controlled online data sharing platform hosted by Swansea University. Only three members of the research team (AT, RC, JR) had access to the data. Their access to the data will cease at the end of this project.

Safe Outputs

All outputs were disclosure checked by an ALSPAC Data Linkage Manager prior to release from UKSeRP.

2. Data Processing Roles

The Data Controller for the information directly collected by ALSPAC is the University of Bristol. ALSPAC is also the (joint) Data Controller for information about participants collected from routine administrative sources.

3. Legal Basis for data processing

ALSPAC's purpose is to conduct scientific research that aims to improve the public good and improve scientific understanding. The legal basis for using participants' information, under GDPR and the Data Protection Act 2018, is:

- 1) performance of a task carried out in the public interest (Article 6(1)(e) in the GDPR); and, where sensitive personal information is involved:
- 2) scientific or historical research purposes or statistical purposes (Article 9(2)(j) in accordance with Article 89(1)).

The GDPR defines ‘sensitive personal information’ as information that reveals a person’s racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership; and the processing of genetic data or biometric data for the purpose of uniquely identifying a person; data concerning health or data concerning sex life or sexual orientation.

This legal basis within GDPR and the Data Protection Act 2018 is separate to, and in addition to, ALSPAC seeking consent to take part in the research process, which they use to help ensure that research is ethical and complies with other applicable laws.

4. Data access process

Access to ALSPAC data is via a managed open access process. Researchers submit their research proposal to the ALSPAC Executive. Once the proposal has been approved, the researchers compile a detailed list of the variables they require, and this is used to build the dataset for the project. AT is an ‘ALSPAC Direct User’ and built the dataset of questionnaire and clinic variables required for this project. This dataset was then passed to an ALSPAC data buddy to check that it matched the approved research proposal, and then the data buddy passed it in turn to an ALSPAC data linkage manager for upload to UKSeRP. The ALSPAC data linkage manager prepared the police and education data required for this project, and placed them, along with the uploaded file, into a specific project folder in UKSeRP. The ALSPAC data linkage manager then performed an ID swap on all the datasets for this project so that the individual ID is unique to this project (meaning the data cannot be linked with any other ALSPAC data). The project folder in UKSeRP can only be accessed by the project team (AT, RC and JR).

3. About the data

3.1. List of variables

Table 3.1: Variable definitions

Variable abbreviation	Variable definition	Variable source	Derivation or specification
Outcomes			
Police recorded serious violence	Any police record for a serious violent crime	A&S Police	This will be derived based on Home Office groupings and age of offence.
Self-reported serious violence	Any self-reported violence at ages 17.5 and 18.5 years	ALSPAC questionnaires	Derived from questions asking if in the past 12 months

			they had: (i) Hit/kicked/punched someone else on purpose with the intention of really hurting them? or (ii) Carried a knife or other weapon for protection or in case it was needed in a fight?
Self-reported weapon carrying	Self-reported weapon carrying at age 15 years	ALSPAC questionnaires	Binary variable derived from the question of whether a participant had in the past 12 months: 'Carried a knife or other weapon for protection or in case it was needed in a fight?'
Exposures			
Bullying victimisation	Any bullying victimisation reported at 8, 10 or 12 years, or at any of these time points; the total number of times reported (0-3).	ALSPAC questionnaires	A participant will be classed as a bullying victim, if they received any of the components of either or both overt or relational bullying frequently (several times a month) or very frequently (several times a week) in the past 6 months.
Bullying perpetration	Any bullying perpetration reported at 8, 10 or 12 years, or at any of these time points; the total	ALSPAC questionnaires	A participant will be classed as a bully, if they had given any of the components of either or both overt or relational bullying frequently (several

	number of times reported (0-3).		times a month) or very frequently (several times a week) in the past 6 months.
Physical bullying involvement	Physical bullying victimisation at any age (8,10 or 12 years). Physical bullying perpetration at any age (8,10 or 12 years).	ALSPAC questionnaires	Derived from a question that pertains to physical bullying victimisation (been hit/beaten up) and perpetration (hit/beaten up someone). Two binary variables will be derived to indicate whether a participant was physically bullied (victimisation) or physically bullied others (perpetration) at any timepoint.
Possible Confounders			
CHILD MEASURES			
Sex	Child sex (male; female)	ALSPAC questionnaires	Standard ALSPAC measure
Age	Age at outcome timepoints (in months)	ALSPAC questionnaires (self-reported violence data); police data (police-recorded violence data)	Standard ALSPAC /police measures
Birth month	Month of the year that child was born	ALSPAC questionnaire data	Standard ALSPAC measure. We may aggregate this to season of birth dependent on numbers.

Ethnicity	Ethnic group (White, non-White)	ALSPAC questionnaires	Standard ALSPAC measure
Behavioural difficulties	Total difficulties score on the Strengths and Difficulties Questionnaire	ALSPAC questionnaire (parent- completed) at age 7	We will either use the raw score or derive a binary indicator of difficulties.
Personality disorder	Childhood Interview for DSM- IV Borderline Personality Disorder (CI- BPD)—11.	ALSPAC clinic at age 11	Standard ALSPAC measure
School enjoyment/attachment	Teacher they can trust, friends at school etc.	ALSPAC questionnaire data.	Variables will be derived from several ALSPAC measures.
SEN	Any special educational needs during KS1 or in Year 3 (yes; no).	NPD data	Dichotomised from SEN variables in the pupil census data
ACEs	Adverse childhood experiences up to 7 years including abuse of child; abuse of mother; parental criminality; maternal mental illness; parental separation/divorce.	ALSPAC questionnaire data.	Measures will be considered individually or as a sum score.
Attainment	Educational attainment at KS1	NPD data (reading, writing and maths test results)	We will derive a categorical variable using quartiles.
Friendships	Friendships	ALSPAC questionnaire (teacher-	Binary indicator of whether child had at least one good friend in the past 6 months.

		reported) at age 7	
FAMILY MEASURES			
Early life SEP	We will consider various measures of early-life socio-economic position (e.g. maternal education, overcrowding, maternal smoking in pregnancy, financial difficulties).	ALSPAC mother questionnaires	These are standard ALSPAC measures, detailed in the data dictionary.
SEP	FSM eligibility (no, yes) during Year 3	NPD data (pupil level census)	Binary measure
Educational engagement	How engaged mother is with child's education	Mother and teacher questionnaires.	These are standard ALSPAC measures, detailed in the data dictionary.
NEIGHBOURHOOD-LEVEL MEASURES			
Residential neighbourhood environment	Deprivation and crime quintiles at end of KS2.	IMD and crime domain from neighbourhood statistics. Income Deprivation Affecting Children Index (IDACI) from NPD.	Quintiles will be derived by linkage data managers.
School neighbourhood environment	Deprivation and crime quintiles at end KS2.	IMD and crime domain from neighbourhood statistics.	Quintiles will be derived by linkage data managers.

3.2. Measurement of key concepts

Table 3.2 Measurement of key concepts

Concept ⁹	How the concept will be measured and encoded
Official sanction for serious violence	The A&SP police data linked to ALSPAC includes records of charges, cautions and other out of court disposals for crimes committed in Avon and Somerset. Serious violence is defined by the Home Office as: (i) violence against the person, indictable only; (ii) robbery, indictable only; and (iii) possession of weapons, triable either way or indictable only. The A&SP data linked to ALSPAC includes Home Office offence codes for each offence.
Self-reported violence	Self-reported violence was defined as saying yes to either or both: (i) Hit/kicked/punched someone else on purpose with the intention of really hurting them? or (ii) Carried a knife or other weapon for protection or in case it was needed in a fight?
Bullying	Bullying will be split into victimisation and perpetration. Both types of bullying involvement were measured when participants were aged 8, 10 and 12 years.
Physical bullying involvement	One of the types of bullying recorded at these timepoints was overt bullying, with a question about physical bullying being given or received (the questions asked if a participant had 'Been hit/beaten up' and if they had 'Hit/beaten up others').

3.3. Missing data and attrition

(i) Key variables that you'd ideally access to address the research question, but will not be present in the datasets

We have very few police records prior to 2007 due to paper records being used before this date, and the fact that the police can only retain records if they have justification to do so (as per Management of Police Information ([MoPI](#)) rules). If we are successful in linking to the PNC, this will mitigate these issues.

When we conduct our preliminary analyses, it may become clear that a small number of the variables we are considering as potential confounders do not have adequate numbers/variability in our sample and so will not be able to be included in our main analyses.

(ii) Observation level data that is incomplete for a subset of the population being studied

As ALSPAC is a longitudinal study, there is attrition and it is known that those from more disadvantaged backgrounds are more likely to drop-out of the study over time. From previous work (not yet published) we know that very few of those with a police record for violence have ALSPAC questionnaire or clinic measures during adolescence. We will include descriptions of our study samples and missing data in both our interim and final reports.

(iii) Potential approaches of mitigating the impact of missing data

We anticipate using multiple imputation to address missing data. We will also take a complete case approach as a sensitivity analyses. We will provide further details on missing data in our interim report.

3.4. Other sources of bias

The school data only include those attending state schools in England. Therefore, children who attend private schools, are home-schooled, or are not registered at a state school in England for any other reason, will not be in our sample.

There are several potential sources of bias with regards the police data. Details are given in the [Data Note](#). In brief, there is known to be bias in terms of whose criminal behaviour is detected by the police, and the disposal type they are given. An example of this is variations in the rate of reporting of crime across communities and demographic groups. With regards the YEF's focus on race equity, it is important to note that ALSPAC is predominately a White UK cohort (>95%), which largely reflects the demographics of the recruitment area at the time the study began in the early 1990s. Therefore, our data are not suitable for examining ethnic differences in the relationship between bullying and violence.

Bias may also be introduced through the data linkage process if participants with a criminal record are, in general, less active in ALSPAC, resulting in their identifier information (e.g. current name and address) held by the study being out of date.

Responses to the questionnaire data may be affected by social desirability bias or recall bias.

Those for whom ALSPAC has permission to link to their education and police records may differ from those who have opted out of linkage or who have not received a consent pack. However, opt out rates for crime and education linkage are low (<4%).

4. About the analysis

4.1. Overview of analytical approach

Our descriptive analyses will include an overview of the dataset, the derivation of our specific study samples for each research question, a summary of each of our exposures and outcomes,

and an examination of potential confounder variables. We acknowledge that it is best practice to specify confounders *a priori*. We have specified the variables that we will consider. However, participation in ALSPAC is known to be sporadic for many individuals (i.e. they may complete a questionnaire or attend a study clinic at one timepoint and then miss one or more before actively participating again). As a result, using several covariates from different timepoints can result in the sample size being dramatically reduced. To avoid this, we may exclude variables if they do not appear to have confounding effects once other factors are taken into account or if numbers with a given characteristic are very small.

We will use logistic regression to model the relationship between our two overall exposure measures (i.e. any bullying perpetration, any victimisation) and our violence outcomes. We will consider the unadjusted association between the exposure (bullying victimisation/perpetration) and outcomes (self-reported violence/official sanction for violence) and then examine the impact of adjusting for a range of early life and childhood factors at the individual, family, and neighbourhood level.

To determine the importance of accumulation of risk, or whether there is a sensitive period, we will use a Structural Lifecourse Modelling Approach (SLCMA) to further examine the relationship between bullying involvement (victimisation/perpetration) and outcomes (self-reported violence/official sanction for violence). This will include those with data from all three timepoints (8, 10 and 12 years).

If numbers allow, we will carry out two additional analyses. Firstly, we will investigate the association between physical bullying (victimisation/perpetration) with the violence outcomes. Secondly, we will study how bullying involvement is associated with weapon carrying at 15 years. [Note that weapon carrying is also included in the main self-reported violence outcome (i.e. at 17.5 or 18.5 years); however, from previous research we know that numbers reporting weapon carrying at these ages are relatively small.]

4.2. Approach to addressing research question(s)

Research questions: approach and methods

Research question	(1) What is the association between bullying perpetration or victimisation and (i) self-reported violent behaviour and (ii) official sanction for serious violence. (2) What is the impact of timing and duration of bullying involvement on (i) self-reported violent behaviour and (ii) official sanction for serious violence.
--------------------------	--

	(3) What is the association between bullying involvement and self-reported weapon carrying in adolescence? (4) Does involvement with physical bullying have a greater associated risk of outcomes than other types of bullying?
<i>Hypothesis, if relevant</i>	Null hypothesis: there is no association between bullying and self-reported violent behaviour/police records for violence.
<i>What will you be able to say by the interim report</i>	For both these research outcomes (self-report and police-recorded violence), we will be able to state the provisional sample size (with flow chart showing derivation of study sample) and highlight how this sample compares to the overall ALSPAC sample in terms of key demographic and confounder variables. We will be able to show the prevalence of bullying (perpetration and victimisation) at each age and the risk of outcomes.
<i>Descriptive analysis, if relevant</i>	For these research questions, In the interim report we will include summary statistics detailing: (1) The provisional size of the study sample for this specific question (2) The main demographic variables for this study sample (3) The prevalence of self-reported and official sanction for violence (4) The prevalence of bullying (also split into subgroups including physical bullying) (5) The extent of missing data
<i>Models, specifications and statistical techniques used, if relevant</i>	We will use logistic regression models for all research questions. We will use a Structured Life Course Modelling Approach (SLCMA) for the analyses of timing and duration of exposure. SLCMA is a statistical method used to examine the association between repeated exposure measures and

	distal outcomes. Using this approach we will examine whether there are sensitive periods of exposure to bullying, during which there is a stronger association between bullying and violence, or whether the number of time points of exposure (duration of exposure) to bullying is more strongly associated with violence, regardless of the time point bullying occurs. This means that as the duration or number of time points of exposure to bullying increases, the risk of violence also increases in a linear pattern. Unlike traditional regression, SLCMA allows us to test both timing and duration hypotheses simultaneously in one model. We will create four variables to test the hypotheses on timing and duration of exposure to bullying: Variable T1 = exposed at age 8 years vs not exposed at age 8 years Variable T2 = exposed at age 10 years vs not exposed at age 10 years Variable T3 = exposed at age 12 years vs not exposed at age 12 years Variable T = T1 + T2 + T3 Variables T1-T3 test timing effects, this will enable us to identify if there are sensitive periods for bullying. Variable T will enable us to test whether duration of exposure to bullying is more strongly associated with violence, irrespective of timing of exposure to bullying. We will present unadjusted and adjusted results.
<i>Estimating equation, if relevant</i>	The included variables will depend on what measures we identify are likely confounders in our descriptive work.
<i>What does the approach need to succeed (constraints/assumptions)?</i>	Variables will only be able to be included if there is adequate variability/numbers in each category. Variables not meeting this requirement will be excluded.

	SLCMA requires that the exposure (bullying) variables are not very highly correlated (defined as $r > 0.90$) as strongly correlated variables diminish the ability of SLCMA to detect the hypothesis best supported by the data. Based on previous research which has used these variables, we do not expect them to be highly correlated.
<i>Uncertainty and inference</i>	Results of the logistic regression models will be given as unadjusted and adjusted odds ratios with 95% confidence intervals and corresponding p-values.
<i>Robustness checks</i>	We will use complete case analyses as a sensitivity check.
<i>Subgroup you intend to study</i>	We do not intend to undertake subgroup analyses but will test for interactions with sex if numbers allow.
<i>Changes to the analysis</i>	As stated, variables will only be included in models where numbers are adequate and when initial analysis indicates that variables are important confounders. The statistician on this project (RC) will be consulted on any other changes to the planned analyses.

5. Project management

5.1. Risks and mitigations

Table 5.1 Risks and mitigations

Number	Risk	Likelihood (Low/Medium/ High)	Mitigation
1	UKSeRP access issues	Very low for long-term issues; medium for one off, short-term issues.	If they occur, these will be reported to UKSeRP and the ALSPAC data linkage team as soon as possible so that can be resolved.

2	Number of participants with both exposure and outcome data is small	Medium	We have chosen to focus our outcome time period for police-recorded violence on ages 13-24 years. We can study police records for any offence if the number with serious violence offences is too small.
---	---	--------	--

5.2. Timeline

Table 5.2 Timeline

Date	Activity	Staff responsible/leading
December 2024	Analysis protocol submission	JR, AT, RC, IB
Oct - Nov 2024	Initial data cleaning	JR (with AT, RC and IB advising)
Jan – Apr 2025	Redrafting of analysis protocol	AT / RC (gap for recruitment of new Senior Research Associate, SRA)
May – June 2025	Data familiarisation and descriptive analyses	SRA
July 2025	Interim first draft report to YEF	SRA (with AT, RC and IB advising)
August 2025	Interim report final draft to YEF	SRA (with AT, RC and IB advising)
July – Oct 2025	Main analysis - analytical models	SRA (with AT, RC and IB advising)
Oct 2025	Final report first draft to YEF	SRA and AT (with RC and IB also contributing).
Dec 2025	Final report final version to YEF	SRA (with AT, RC and IB advising)

6. References

1. Lasher S and Baker C. Bullying: evidence from the Longitudinal Study of Young People in England 2: Research Brief 2015. Available from: https://assets.publishing.service.gov.uk/media/5a7f2ccced915d74e33f4cb5/Bullying_evidence_from_the_longitudinal_study_of_young_people_in_England_2__wave_2_brief.pdf.
2. Stassen Berger K. Update on bullying at school: Science forgotten? *Dev Review* 2007; 27(1):90-126.
3. Walters GD. School-age bullying victimization and perpetration: a meta-analysis of prospective studies and research. *Trauma, Violence, Abuse* 2021; 22(5):1129-1139.
4. Armitage R. Bullying in children: impact on child health. *BMJ Paediatr Open* 2021; 5(1): e000939.
5. Ye Z et al. Meta-analysis of the relationship between bullying and depressive symptoms in children and adolescents. *BMC Psychiatry* 2023; 23(1):215.
6. Vrijen C et al. Childhood and adolescent bullying perpetration and later substance use: a meta-analysis. *Pediatrics*, 2021. 147(3).
7. Stapinski LA et al. Peer victimization during adolescence and risk for anxiety disorders in adulthood: a prospective cohort study. *Depress Anx* 2014; 31(7):574-582.
8. McVie S. The impact of bullying perpetration and victimization on later violence and psychological distress: a study of resilience among a Scottish youth cohort. *J Sch Violence* 2014; 13(1):39-58.
9. Laith R and Vaillancourt T. The temporal sequence of bullying victimization, academic achievement, and school attendance: A review of the literature. *Aggress Violent Behav* 2022; 64:101722.
10. van Geel M, Vedder P, and Tanilon J, Bullying and weapon carrying: a meta-analysis. *JAMA Pediatrics* 2014; 168(8):714-720.
11. Valdebenito S, Ttofi MM, Eisner M, Gaffney H. Weapon carrying in and out of school among pure bullies, pure victims and bully-victims: a systematic review and meta-analysis of cross-sectional and longitudinal studies. *Aggress Violent Behav* 2017; 33:62-77.
12. Ttofi MM, Farrington DP and Lösel F. School bullying as a predictor of violence later in life: a systematic review and meta-analysis of prospective longitudinal studies. *Aggress Violent Behav* 2012; 17(5):405-418.
13. Farrington DP and Ttofi MM. Bullying as a predictor of offending, violence and later life outcomes. *Crim Behav Ment Health* 2011; 21(2):90-8.
14. Tiiri E et al. Bullying at 8 years and violent offenses by 31 years: the Finnish nationwide 1981 birth cohort study. *Eur Child Adolesc Psychiatr* 2023; 32(9):1667-1678.
15. Brennan IR and Moore SC. Weapons and violence: a review of theory and research. *Aggress Violent Behav* 2009; 14(3):215-225.

16. House of Common Library. Knife crime statistics: England and Wales. House of Commons Library 2023.
17. Vulliamy P et al. Temporal and geographic patterns of stab injuries in young people: a retrospective cohort study from a UK major trauma centre. *BMJ Open* 2018; 8:e023114.
18. Ttofi MM and Farrington D. School bullying: risk factors, theories and interventions. In Brookman F, Maguire M, Pierpoint H and Bennett T (eds). *Handbook on Crime* (pp.427-457). Willan 2010.
19. Lucas S et al. Bully, bullied and abused. Associations between violence at home and bullying in childhood. *Scand J Public Health* 2016; 44(1):27-35.