

Sure Start – Measuring and improving its effectiveness (NESS)

Ted Melhuish

2000

Tender to evaluate Sure Start

Randomised Control Trial ruled out as a methodology

2001 – National Evaluation of Sure Start

4 major components

1. Local Context Analysis – study of communities over time

2. Implementation – what programmes do

3. Impact – do programmes affect children and families

4. Cost-effectiveness – how money is spent – and is it effective

We will concentrate here on the Impact Study

Impact study

Aim:

To evaluate impact of SSLPs upon children and families

Three components to Impact Study

First phase – cross-sectional study

- 9 and 36 month olds and their families
- in SSLP and SSLP-to-be areas

Programme variability

- investigates links between implementation and impact

Longitudinal study

- 8000 children seen at 9 months, 3 years, 5 years
- comparison group from Millennium Cohort Study

As RCT ruled out we had to consider a quasi-experimental design.

Who should be the control group?

-the answer to this had to change as the study progressed.

Cross-sectional phase, controls=Sure Start-to-be -waiting list controls

Longitudinal Phase – Sure Start increased in size and Sure Start-to-be now receiving Sure Start so had to look elsewhere

– Millennium Cohort Study (MCS)

Missing data – bias?

We can impute values for missing data – using multiple imputation techniques (Rubin 1997).

So we have 2 datasets

1. Complete cases data – some cases left out if missing data

2. Imputed data – missing data imputed so all cases in analysis

Dilemma

- complete cases analyses may have bias

But

- we cannot be absolutely certain of validity of imputation

Solution

Only use results that are significant and consistent in both complete cases and imputed datasets.

- Conservative strategy

INTENTION TO TREAT DESIGN

All children and families in an area are the targets of Sure Start
so random sample in an area chosen

Sample chosen from Child Benefit records

2005: Cross-sectional results

Sub-group findings (3-year-olds)

Among non-teenage mothers (86% of total):

- greater child social competence in SSLP areas
- fewer child behaviour problems in SSLP areas
- less negative parenting in SSLP areas

Effects on children appeared to be mediated by effects on mother:

- SSLP → less negative parenting → better child social functioning

2005: Sub-group findings (3-year-olds)

Among teenage mothers (14% of total):

- less child social competence in SSLP areas
- more child behaviour problems in SSLP areas
 - poorer child verbal ability in SSLP areas

Among lone parent families (40% of total):

- poorer child verbal ability in SSLP areas

Among children living in workless households
(33% of total):

- poorer child verbal ability in SSLP areas

Also large variation amongst SSLPs

Key question:

Why are some SSLPs more effective in achieving outcomes than others?

Programme variability provides some answers

Impact study uses multi-level modelling

Data clustered by Sure Start areas

Therefore can create measures of “effectiveness” from the multi-level model residuals at the area level, having allowed for all other covariates (large number of background characteristics)

Effective SSLPs have outcomes better than expected based on covariates

Ineffective SSLPs have outcomes worse than expected based on covariates

Therefore we have a continuum of “effectiveness”

Programme variability

What predicts “effectiveness” for child and family outcomes?

We use all our data on programme implementation to construct 18 dimensions of proficiency.

Key dimensions of proficiency of implementation for effectiveness:

Effective governance and management / leadership

Informal but professional ethos of centre

Empowerment of service providers and users

Programme variability

Key strategies to attain proficiency:

Tuning into local community for universal services

Early identification and treatment for specialist services

Recruiting / training staff – qualifications & attitudes

Managing multi-agency teamwork

Programme variability

- Challenges for children's centres and training:
- Higher reach needed (especially overcoming barriers for 'hard to reach')
- Better multi-agency teamwork
- Sustainable, shared systems for monitoring service use / treatments
- More rigour in measuring impact of treatments
- Grasp of cost effective deployment of specialist / generalist staff
- Co-ordinating outreach and centre based services

Current impact report - 2008



The impact of well-established SSLPs on
3-year-olds and their families

How to find comparison group?

Millennium cohort study – random sample of children in UK
Using post codes for sample in England

- Create clusters resembling neighbourhoods
- Select MCS neighbourhoods not receiving Sure Start
- Using propensity matching on 85 area variables find areas that resemble Sure Start areas
- Use the MCS sample in those areas as comparison group

“propensity scoring”, -addresses selection bias



—that is, the possibility that those who experience a treatment (i.e., Sure Start) may differ in unmeasured ways from those who did not. The term propensity refers to “a conditional probability of an individual being in a treatment group, given a set of background variables for that individual”.

In this study whether a child is in the treatment group is determined by whether or not the child lives in a SSLP area; the problem therefore reduces to identifying those areas that have a greater or lesser propensity of having populations that are similar to those of SSLP areas.

In propensity matching we used **85** area-level variables measured for Sure Start and non-Sure Start areas derived from the 2004 Index of Multiple Deprivation (IMD) and 2001 Census

Propensity scoring used to estimate area similarity, and create “treatment” and “control” groups matched on their propensity to be a SSLP area.

As long as important variables not omitted, the comparison between SSLP and MCS groups should have minimal bias due to the non-random allocation of SSLPs to areas.

SSLP populations more disadvantaged than MCS. This necessitated dividing the NESS and MCS samples into five strata reflecting the degree of propensity to be chosen as an SSLP area.

Stratum 1 - lowest propensity to be chosen as SSLP area

Stratum 5 - highest propensity to be chosen as a SSLP area.

Propensity	SSLP	<i>MCS</i>	<i>MCS</i>
Stratum	N areas	<i>N areas</i>	<i>N children</i>
1	2	<i>53</i>	<i>1041</i>
2	15	<i>40</i>	<i>970</i>
3	33	<i>22</i>	<i>818</i>
4	45	<i>10</i>	<i>565</i>
5	55	<i>1</i>	<i>21</i>
TOTAL	150	<i>126</i>	<i>3415</i>

To accommodate different distributions by strata a two-stage analysis plan was implemented

1. We restricted the main Sure Start- non-Sure Start comparisons to Strata 2-4.
2. Are children/families in strata 2, 3, 4 and 5 functioning similarly.

If they scored similarly on outcome measures, this would suggest that any detected effects of SSLPs should generalise to all Sure Start children/families.

The impact of well-established SSLPs on 3-year-olds & their families

- 5883 children / families in 93 SSLP areas, and
- 1879 children / families in 72 non-SSLP (MCS) areas

Methodology

Using data from NESS and MCS matching areas that were carefully chosen we can compare –

14 outcomes in common for NESS and MCS at 3 years

Results

Controlling for child, family and area characteristics we test for SSLP vs. non-SSLP differences

Of 14 outcomes 7 showed a significant difference between SSLP and non-SSLP areas, i.e. a SSLP effect

Results

Of 14 outcomes 7 showed a significant difference between SSLP and non-SSLP areas, i.e. a SSLP effect

5 outcomes clearly indicated beneficial effects for SSLPs. These were for:

- child positive social behaviour (cooperation, sharing, empathy)
 - child independence / self-regulation
(works things out for self, perseverance, self-control)
- Parenting Risk Index (observer rating + parent-child relationship, harsh discipline, home chaos)
 - home learning environment
 - total service use

In addition there were better results in SSLPs for:

- child immunisations
- child accidents

But these 2 outcomes could have been influenced by timing effects

Methodological Issues



Timing – 2 year gap in data collection between Sure Start and comparison data – we tested for possible effects of timing

Different teams collecting data – we coordinated with MCS in planning data collection and training their staff.

Unmeasured variables – always a problem with quasi-experimental studies (but also RCTs) – we had large number of covariates to reduce the likelihood of unmeasured effects.

What does this mean for the future?

These benefits in terms of parenting and child development have a good prognosis. All are desirable effects that are likely to lead to better long-term outcomes for children

In particular we have good evidence (e.g. from EPPE) that higher child independence and higher HLE are likely to lead to better long-term outcomes both intellectually and socially for children. The other beneficial outcomes support this view

Outcomes with no SSLP effect

There were no differences between SSLP and non-SSLP areas for:

Mother's	Body Mass Index smoking life satisfaction rating of area
Father's	involvement with child
Child	language development negative social behaviour

Do SSLP effects vary by subgroups?

We looked at subgroups by 6 demographics

- gender
- ethnic group
- teen / not teen mother
- lone parents
- workless households
- income (below poverty line or not)

We concluded that the SSLP effects do not vary substantially for the different sub-populations

Do Sure Start areas included (strata 2,3,4) differ from those **not** included in comparisons (strata 5).

We analysed for significant differences in models of the 14 outcomes between these 2 groups.

The models applied equally well to all Sure Start areas i.e. similar child and family functioning in Sure Start areas across strata

Why are results now so different to the earlier report?

We need to acknowledge that there are methodological differences between the first phase and the current phase of the NESS impact study

However there are good substantial reasons for why the results are different now

Reasons for differing results

1. Amount of exposure

It takes 3 years for a programme to be fully functional. Therefore

a. in the first phase children / families were not exposed to fully functional programmes for much of the child's life

b. in the second phase children / families are exposed to fully functional programmes for all child's life

2. Quality of services

a. SSLPs have been reorganised as SSCCs with clearer focus to services following lessons from earlier years, and from NESS

b. early on staff had a lot to learn. As knowledge and experience have been acquired over 7 years, SSLPs have matured in functioning and staff skill shortages have reduced

c. hence it is likely that children / families are currently exposed to more effective services than in the early years of Sure Start

Conclusion

The impact of Sure Start has improved, probably because of:

1. increasing quality of service provision, greater attention to the hard to reach, the move to children's centres

as well as

2. the greater exposure of children and families

These positive results are modest but are evidence that the impact of Sure Start programmes is improving

Overarching messages

1. Programmes have improved over the years and Children's Centres are in the right direction
2. Many examples of good practice
3. There is still great variation between best and worst
4. Need to learn from most effective Children's Centres

Overarching messages – cont.

5. Inter-agency collaboration is essential for good services
6. Active engagement of health services important for success of Sure Start. Health has contact with all families and children from pregnancy
7. However beneficial services are, children and families need to be in touch with them
8. Those with the greatest need may be hardest to reach and engage

Overarching messages – cont.

9. Trust is fundamental to parental engagement
10. Staff capacity problems, many staff
inadequately trained for the work to be
done and
staff turnover is very disruptive

My personal choice for top priority

**Need to increase focus on
child language development**

NESS website address

www.ness.bbk.ac.uk