The Apartheid regime created a system of so-called race classification which characterized ‘the struggle’. Black Africans began to disregard restrictive legislation that constrained where they lived and worked. Very rapid unplanned urbanization began, and shanty towns mushroomed around formerly White cities and towns. It was anticipated that this rapid urbanization, with urban growth estimated at the time to be 3.5% per year, would have profound effects on children’s health and development. While improved access to health care, education and employment in urban areas could decrease preventable childhood morbidity and mortality, the inability of government to establish and maintain services to meet the needs of the growing urban population could exacerbate existing infectious diseases, such as HIV/AIDS and tuberculosis. Non-infectious conditions related to the interaction of lifestyle, urban stressors and socio-cultural changes—for example, childhood injuries, substance use and obesity—were predicted to increase.

In 1988, as a result of discussions on these issues, Noel Cameron at the University of the Witwatersrand, and Derek Yach at the South African Medical Research Council (MRC) approached the then MRC President, Andries Brink, for funds to begin a birth cohort study in Soweto-Johannesburg with Lucy Wagstaff, also from the University of the Witwatersrand. The aim was to track a group of urban children for 10 years, not knowing at the time, that these children would also be the first cohort born into a democratic South Africa. At the start, the study was called Birth to Ten (BT10), but changed to Birth to Twenty (BT20) in 2000 when we reached the 10-year follow-up goal and resolved to continue the study to age 20. The children came to be known colloquially as Mandela’s Children because they were born in the 7 weeks following Nelson Mandela’s release from prison on the February 11, 1990, an event that heralded in radical social and political changes. Linda Richter was part of the original investigator group. Shane Norris has been the Project Manager since 2001.

With seed funding, we undertook several pilot studies to determine, amongst other things, the seasonality of births in the area and the optimum months for recruitment into a birth cohort study; the nature and accuracy of routinely collected health service data; and follow-up rates of children from birth to six months. The date for enrolling the birth cohort was set for February, but a national hospital strike delayed enrolment, which occurred from April 23 to June 8, 1990.

What does the study cover—and how has this changed?

From its inception, BT20 was planned to be multidisciplinary, tracking the growth, health, well-being and educational progress of urban children across the first decade of their life. To do so, we had to innovate in a number of areas, such as establishing tracking systems under circumstances where few people had a street address or telephone, create a flexible dataset because names were inaccurately translated and transcribed from record to record, translating and adapting questionnaires developed in the West and establishing norms for many of the measures used.

We began the study during the antenatal period, collecting information on pregnancy and birth. Because we adopted a lifecycle approach, the study covers many of the major issues confronting the particular developmental phase of children and young people in which data is collected. In the early years we concentrated on environmental influences (poverty, migration and political violence), access to health services, nutrition, child care and growth and development.

In the first few years of school, the emphasis was on cognitive ability and school performance, as well as social adjustment

---

1 Human Sciences Research Council and University of KwaZulu-Natal, South Africa.
2 University of the Witwatersrand, South Africa.
3 Rockefeller Foundation, USA.
4 Loughborough University, UK.
5 The Apartheid regime created a system of so-called race classification which legally differentiated between Whites (of European origin), Indians (a collection of different people from the South East Asian region, mainly from India), Coloureds (people of mixed ancestry) and Blacks (people of African descent). Participation in society was differentiated in freedom and quality on a continuum from Whites to Indians to Coloureds to Black people. The terminology is retained because it carries the legacy of decades of oppression and discrimination, the effects of which are still evident.
6 Corresponding author. Child Youth Family and Social Development, Human Sciences Research Council, Private Bag X07, Dalbridge 4014, South Africa. E-mail: lrichter@hsrc.ac.za
and participation. Later questionnaires began to explore early lifestyle risk factors—diet and weight gain, parental monitoring and supervision, educational failure, and sexual experimentation, and a wide range of physical (body composition and bone mineral density scans) and physiological measures (pubertal development, biochemical markers of insulin resistance) were introduced. Now that the cohort has turned 16, data collection is focused on the prediction and measurement of risk, particularly in three domains: (i) sexual and reproductive behaviour (sexual debut, sexual violence, unplanned pregnancy and sexually transmitted infections), (ii) early expression of the metabolic syndrome (obesity, hypertension and insulin resistance) and (iii) social marginalization (school drop out, psychosocial difficulties, substance abuse, violence and conflict with the law).

The major change in the study occurred as we neared our 10-year goal. BT10, as it was then called, was a collaborative enterprise involving more than 20 researchers, some of whom had interests in only one specialized area or particular age group. No single individual had responsibility for the overall direction and resourcing of the study; there was a constant struggle to make ends meet and to enter and clean data requested by individual scientists. In response to this situation, Linda Richter and John Pettifor took the decision to raise research funding for two specific hypothesis-driven longitudinal studies, one on sexual risk and the other on bone health. BT10, as it had existed, was closed down and BT20 took its place, incorporating all the staff and data from the first ten years.

As from October 2005, the next generation of children has started to be born.1 Our first young mother was 14-years-old when she delivered her baby. There were three 14-year-old mothers in the original cohort of parents. Deliveries are speeding up; more than a third of women in South Africa have their first child before their 20th birthday. While we intend to repeat sampling up—and what is attrition like?

Table 2 shows the year, place of data collection, numbers contacted, from whom data was collected and reasons for attrition at all data collection waves to date.

We distinguish three types of attrition: absolute attrition due, for example, to the death of a child. To determine this, we have relied on reports of family and neighbours when visiting the last known address of the family. Second, there is intermittent attrition due to circular migration between rural and urban areas. This is evident in the large numbers of people (n = 1449) who are reported to have left the study area during any particular data collection wave (in total 44% of the cohort), some of whom return to the urban area and are traced during a later data collection wave. Last, there is study attrition, due to the incapacity of the study to either trace the family or interview the participants in a data collection wave, because of a lack of resources, including time. For example, the 9–10-year data collection wave was never completed. A proportionately larger number of families were reported to have moved during the data collection years 3 and 4 because we devoted a large amount of time to tracing families who had not been found during the previous few years.

**Who is in the sample?**

BT20 includes all singleton children born to women resident in Soweto-Johannesburg during the 7-week enrolment period. In late 1989, we began interviewing women in public antenatal clinics in an effort to collect data from pregnant women who were predicted to deliver their babies during the enrolment period. More than 2000 women participated; however, because the cohort enrolment dates shifted as a result of the hospital strike, only 1594 of these women delivered within the enrolment period. A further selection criterion was that the mother and baby had to remain in the area until the child was six months old. The reason for this criterion was that the pilot studies had demonstrated that large numbers of women came from rural areas to deliver babies in the city, either because they believed the services were better, to ensure their child a birth certificate entitling them to live and work in an urban area under Apartheid law, or because they wanted the child’s father, a migrant labourer, to witness his child’s birth and thus secure the child’s support. All births are registered in the municipal area under a local ordinance, but to ensure that our records were complete, we also checked mortuaries and back-tracked children brought to clinics for their first 6-week postnatal check-up. According to the records we compiled, 5449 births were registered during the 7-week enrolment period, 3273 of which met the cohort entry criterion of continued residence in the area. Of this group, we are still in contact with more than 70% of the cohort at age of 16 years.

Birth notification forms contain basic information from which we could derive characteristics of total births. In terms of so-called race groups, the total cohort is roughly representative of the South African population, except for an under-representation of White children. Whites comprise 9.2% of the population of the country.5 The reason for their under-representation is 2-fold. First, it was not possible to enrol children through private health practitioners and facilities, which White families predominantly used at the time. As the study has progressed, White families, who tend to be better off, have shown higher attrition than other families. Cohort studies in developing countries have generally found, contrary to what pertains in better-resourced countries, that higher socioeconomic class individuals are harder to enrol and maintain, principally because they perceive fewer benefits and more disadvantages to participation in research.6 To compensate for this, at age 10 years in the bone health study, we recruited a supplementary sample of 120 White children born during the cohort enrolment dates, but not in the area. Most migrated into the metropole between 1990 and 1997. Table 1 shows some basic characteristics of the cohort.7,8 Although retrospective data has been collected from the supplementary group, they are not included in the description of the original cohort in the table.

How often have they been followed up—and what is attrition like?

Table 2 shows the year, place of data collection, numbers contacted, from whom data was collected and reasons for attrition at all data collection waves to date.

We distinguish three types of attrition: absolute attrition due, for example, to the death of a child. To determine this, we have relied on reports of family and neighbours when visiting the last known address of the family. Second, there is intermittent attrition due to circular migration between rural and urban areas. This is evident in the large numbers of people (n = 1449) who are reported to have left the study area during any particular data collection wave (in total 44% of the cohort), some of whom return to the urban area and are traced during a later data collection wave. Last, there is study attrition, due to the incapacity of the study to either trace the family or interview the participants in a data collection wave, because of a lack of resources, including time. For example, the 9–10-year data collection wave was never completed. A proportionately larger number of families were reported to have moved during the data collection years 3 and 4 because we devoted a large amount of time to tracing families who had not been found during the previous few years.

---

1 Called 3G because they are the third generation; 3G also refers to a high-speed wireless internet connection, a fitting association for these children, born to parents half of whom have a cell phone.
Follow-up in the early years was severely handicapped by a lack of resources. At the time, we restricted follow-up of children to the original cohort area, some 400 km². With the collapse of Apartheid, the urban landscape changed substantially, because families had the freedom to live and work where they chose (Figure 1). Currently, we trace children and families throughout the Gauteng province (an area of 17 000 km²).

In 1993 (cohort age 3), we suspended data collection in order to devote all our capacity to tracing ‘lost’ children and families, and in 2003 (cohort age 13), we conducted a cross-sectional survey of same-aged children in 622 primary schools in the area, primarily to trace families lost to follow-up. Through the school survey, we re-connected with close to 400 children previously thought to be lost because of multiple address changes. This is a generally very mobile population, with about a quarter of families reporting a change of address during each major data collection period. There are 441 families who we saw once within the first year of life but not again. These families probably moved and are now considered absolute losses to follow-up.

We use multiple strategies to maintain participation. When the study began, we created relationships with the community and early data collection was planned in cooperation with the health services. A Community Advisory Board (CAB) was created whose membership has changed over time to reflect current concerns in the study. Currently, education officials and parents are strongly represented. In addition, field workers were recruited from local communities, some of whom have been employed in the study from its inception. Staff is also trained in cohort maintenance, as well as Good Clinical Practice and ethical research conduct.

The maintenance strategy, especially in the first decade, was directed to parents and other primary caregivers. Dissemination of information about the study was crucial to ensure that the requirements were well understood and participants were motivated to take part. Information is disseminated by staff at each contact, as well as through newsletters and workshops. Compensation for transport expenses is provided for families who come into one of the two data collection sites, at the rate of R50.00 a visit. Tracing of families is done through multiple contact addresses and telephone numbers collected from all participants at each contact.

Contact with families is maintained between data collection phases through birthday cards, reports on the study in newspapers and on radio and television and regular newsletters for parents as well as for child/adolescent participants and school principals. A limited social and health referral service is incorporated into the study—for example, a toll-free telephone number and a free-post address to enable families to contact the office for advice and information; and referral notes to local services are given to families when health or social problems are detected.

As the children have grown older, the cohort maintenance strategy has been re-orientated to the adolescent participants. On average, study participants spend 3–4 h at the data collection site, during which time refreshments are provided. The study team provides letters to schools explaining children’s absence if they miss school. We also promote awareness of BT20 at schools. However, the most significant initiative has been to shift data collection to weekends to make it more convenient for caregivers and adolescents to visit the site.

### What has been measured?

While some rounds of data collection contain questionnaires or measures of special interest, common themes run through the data collection rounds, as shown in Table 3. These include demographic, socioeconomic and household information; community, neighbourhood and school environments; health and nutrition; childcare, supervision and monitoring; growth and physical activity; cognitive development and school performance; social and psychological adjustment; risk behaviours; and a range of physiological measures. Data has been collected at 17 points across the 16 years of the study—at antenatal delivery, 6 months, 1 year, 2 years, 3 years, 4 years, 5 years, 7 years, 9/10 years, 11/12 years, 13 years twice in the year, 14 years twice and 15 years twice—as well as through a school survey at age 13 years.

### What has been found?

A full list of publications is available on the Birth to Twenty website (http://www.wits.ac.za/birthto20). Under the categories

---

§ Exchange rate $1=R7.
used on the web site, highlights from each of five areas are briefly summarized as follows.

**Methodological issues**

At the start of the study, several analyses examined the reliability and validity of routinely collected health service data, including reproductive histories reported in response to questionnaires\(^{10}\) and the accuracy of birth data recorded on health and administrative records.\(^{11}\) The information collected was used to design subsequent questionnaires. Several papers engaged vigorously with the issue of race in the new South Africa,\(^{12}\) as well as implications for care. For example, we found that White women using private facilities were ten times more likely to have a Caesarean section delivery than African women using a public health service.\(^{13}\) More recently, methodological substudies have concentrated on ensuring accurate measurement of pubertal status\(^{14}\) and maximizing disclosure of sensitive information by young adolescents.\(^{15}\)

**Nutrition**

From longitudinal food frequency analyses—conducted in 1995, 1997, 1999 and 2000—we consistently found that calcium, iron, zinc and biotin fell below 65% of the Recommended Daily Allowance. In addition to these nutrients, vitamin A, riboflavin, nicotinic acid and pantothenic acid fell below the RDA in 2003. From 2000 to 2003 there was an increase in the percentage of children falling below the recommended intakes for most micronutrients.\(^{16,17}\)

**Growth and bone health**

The cohort shows improved growth in comparison with children born in the 1970s, but with little reduction in the difference between Blacks and Whites. White children continue to be taller and heavier.\(^{18}\) Analyses examining ethnic differences in bone mass indicate that pre-pubertal Black children, despite adverse environmental factors (low calcium intakes <400 mg/day, less physical activity, lower socio-economic status) have greater hip bone mass than White children, and similar bone mass at the spine and appendicular skeleton after adjusting for differences in stature, likely due to genetic rather than environmental factors.\(^{19}\) Physical activity, as measured by the amount of sport played, walking, and leisure activities, is lower in Black than White children, mainly because of poor physical education and sporting opportunities in traditionally Black schools.\(^{20}\) The amount of physical activity undertaken by Black children is insufficient to have an osteogenic effect on bone mass. Peak bone mass is established in childhood and this influences the likelihood of minimal trauma fractures in later life. A lack of physical activity in childhood may also contribute to obesity in young adults.

**Infant, child and adolescent well-being**

Similar to their counterparts in resource-rich countries, young children in South Africa show age-related increases and decreases in adjustment profiles, gender differences in externalizing and internalizing behaviours, and early onset and persistence of maladjustment amongst a small group of children whose first signs of difficulties manifest as

---

**Table 2** Number of study participants contacted and lost to follow-up over the past 10 data collection waves

<table>
<thead>
<tr>
<th>Year reported in this data wave</th>
<th>Antenatal 6 months</th>
<th>1 year</th>
<th>2 years</th>
<th>3/4 years</th>
<th>5 years</th>
<th>7/8 years</th>
<th>9/10 years</th>
<th>11/12 years</th>
<th>13 years</th>
<th>14 years</th>
<th>15 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place of data collection</td>
<td>Clinic &amp; Homes</td>
<td>Clinic &amp; Homes</td>
<td>Clinic &amp; Homes</td>
<td>Clinic &amp; Homes</td>
<td>Clinic &amp; Homes</td>
<td>Clinic &amp; Homes</td>
<td>Clinic &amp; Homes</td>
<td>Clinic &amp; Homes</td>
<td>Clinic &amp; Homes</td>
<td>Clinic &amp; Homes</td>
<td>Clinic &amp; Homes</td>
</tr>
<tr>
<td>Established contact</td>
<td>20 (1.0%)</td>
<td>3 (0.0%)</td>
<td>9 (0.4%)</td>
<td>16 (0.8%)</td>
<td>16 (0.8%)</td>
<td>16 (0.8%)</td>
<td>16 (0.8%)</td>
<td>16 (0.8%)</td>
<td>16 (0.8%)</td>
<td>16 (0.8%)</td>
<td>16 (0.8%)</td>
</tr>
<tr>
<td>Reported child or caregiver mortality</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Family moved out of the study area</td>
<td>46 (5.0%)</td>
<td>56 (5.6%)</td>
<td>199 (19.8%)</td>
<td>298 (29.8%)</td>
<td>288 (28.8%)</td>
<td>182 (18.2%)</td>
<td>182 (18.2%)</td>
<td>182 (18.2%)</td>
<td>182 (18.2%)</td>
<td>182 (18.2%)</td>
<td>182 (18.2%)</td>
</tr>
<tr>
<td>Family moved to unknown address</td>
<td>39 (2.5%)</td>
<td>51 (2.5%)</td>
<td>218 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
</tr>
<tr>
<td>Total</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
</tr>
<tr>
<td>Percentage (%) of the cohort lost to follow-up in this wave</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
<td>327 (2.5%)</td>
</tr>
</tbody>
</table>
parental perceived ‘difficultness’ and poor peer relations. Exposure to violence, in the home and the community, has been found to have a pervasive relationship with children’s emotional and social adjustment, independent of the impact of poverty.

Sexual and lifestyle risk

Sexual risk
Unwanted teen pregnancy and a high prevalence of sexually transmitted infections (STI), including HIV, are significant health threats in South Africa, both increasing as BT20 children move into their late teens. To date, 15 girls in the cohort have delivered a child, the first just after turning 14 and a further 13 are pregnant. Sexual intercourse was reported among 5% of 13-year-olds, equally by boys and girls; it picks up rapidly at 14 (17% boys, 6% girls) and 15 (27% boys, 11% girls). However, amongst both boys and girls more than half of all sexual activity reported, including foreplay and oral sex, is indicated to be non-consensual. Two BT20 adolescents have tested HIV positive through a voluntary screening programme and several have been treated for other sexually transmitted infections.

Metabolic disease risk
Children who are born small and then grow quickly appear to be at increased risk of obesity and risk factors for Type II diabetes. We found that lower birth weight was associated with greater insulin production and higher glucose concentrations at age 7. This suggests that the ability of insulin to clear glucose from the circulation is reduced in children with low birth weights. Body mass index at 7 years was positively associated with both insulin concentrations and insulin resistance, as was the rate of weight change or weight velocity from birth to 7 years. Finally, high subcutaneous fat levels at the subscapular and triceps sites, measured at age 7, were also positively related to both insulin concentrations and insulin resistance.

What are the main strengths and weaknesses of the study?
A critical strength has been the perseverance of a small group of people determined to maintain the study, with little funding and against some opposition. At the start, several prominent public health specialists warned that long-term follow-up was neither possible nor particularly useful. The core group, who also provide continuity, drove a collaborative, multidisciplinary approach and a strong association with health services, schools and the community.

As resources have grown, most significantly with long-term funding from the Wellcome Trust, stable systems have been established for the management of the study. A team of very committed people, responsible for operations, administration, laboratory and data, report to the investigators on a weekly basis. Sophisticated bar code, filing, and electronic systems have been designed to print address lists, weekly appointments, tracking participants through the study components, data...
Table 3  Data collected in Bt20

<table>
<thead>
<tr>
<th>General category</th>
<th>Example</th>
<th>Data collection wave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residence and contact details</td>
<td>Address, relatives/friends contact numbers</td>
<td>All</td>
</tr>
<tr>
<td>Growth</td>
<td>Birth weight, height, weight, skinfold circumferences</td>
<td>All</td>
</tr>
<tr>
<td></td>
<td>Head circumference</td>
<td>3 months to 7/8 years</td>
</tr>
<tr>
<td></td>
<td>Hip and waist circumferences</td>
<td>From 9 years onwards</td>
</tr>
<tr>
<td>Socio-economic circumstances</td>
<td>Education; occupation; housing type and ownership; water, sanitation &amp;</td>
<td>0–2 years; 5 years; 7/8 years; 11/12 years; 13/14 years</td>
</tr>
<tr>
<td></td>
<td>electricity; refuse removal; household composition and density; household assets</td>
<td></td>
</tr>
<tr>
<td>Child health</td>
<td>Morbidity; health histories; immunization; accidents; injuries; developmental problems; medication; accessing health services</td>
<td>6 months; 1 year; 2 years; 4 years; 7/8 years; 11/12 years; 13 years; 14 years; 15 years</td>
</tr>
<tr>
<td>Maternal/carer/giver stress and health</td>
<td>Stress and social support; post-natal depression; family medical history; alcohol and tobacco use; lifestyle; weight; height; body composition; blood pressure</td>
<td>Antenatal: 6 months, 5 years, 7 years, 10 years, 11/12 years, 13 years, 14 years, 15 years</td>
</tr>
<tr>
<td>Child nutrition</td>
<td>Breastfeeding practices, dietary intake (food frequency) retrospective recall</td>
<td>0–2 years; 5 years; 7/8 years; 9 years; 10 years; 13 years</td>
</tr>
<tr>
<td>Child care</td>
<td>Caregiving; day care; after school activities; monitoring and supervision; perception of parenting and family relationships</td>
<td>0–2 years; 5 years; 7/8 years; 11/12 years; 13–15 years</td>
</tr>
<tr>
<td>Cognitive development</td>
<td>Pre-speech development; Bayley Scales of Infant Development; Griffiths Scales of Mental Development; speech, language and hearing development; Denver Developmental Screeing Questionnaire; Ravens Coloured Progressive Matrices</td>
<td>6 months, 1 year, 2 years, 4 years, 5 years, 7 years</td>
</tr>
<tr>
<td>Education</td>
<td>Child care; preschool attendance; school start age; repetitions and failure; annual school reports; teacher ratings of children; school infrastructure; bullying; perceptions of school environment</td>
<td>7/8 years onwards</td>
</tr>
<tr>
<td>Psychological assessment</td>
<td>Richman &amp; Graham Behaviour Rating; South African Child Assessment Scale (SACAS); Draw-A-Person; Youth Self-Report (YSR); Rosenberg Self-Esteem; body image; future aspirations; eating attitudes; identity</td>
<td>2 years, 4 years, 5 years, 7 years, 10 years, 11/12 years, 13 years, 14 years, 15 years</td>
</tr>
<tr>
<td>Neighborhood, community</td>
<td>Community rating; perceptions of safety and connectedness; community violence</td>
<td>5 years, 11/12 years, 13 years, 14 years, 15 years</td>
</tr>
<tr>
<td>Child blood pressure</td>
<td>Diastolic, systolic, pulse rate</td>
<td>5 years, 7 years, 11/12 years, 13 years, 14 years, 15 years</td>
</tr>
<tr>
<td>Physical activity</td>
<td>Transport to school; participation in school sports; sedentary behaviour; extra-mural activities; casual work</td>
<td>7 years, 11/12 years, 13 years, 14 years, 15 years</td>
</tr>
<tr>
<td>Bone mass &amp; body composition</td>
<td>DXA-derived bone mass and body composition</td>
<td>9 years, 10 years, 11 years, 12 years, 13 years, 14 years; 15 years</td>
</tr>
<tr>
<td>Pubertal development</td>
<td>Menstruation; Tanner pubertal development scales</td>
<td>Annually from 11 years</td>
</tr>
<tr>
<td>Biochemical tests</td>
<td>Lead</td>
<td>At birth</td>
</tr>
<tr>
<td></td>
<td>Lipids</td>
<td>5 years; 13 years; 15 years</td>
</tr>
<tr>
<td></td>
<td>Glucose and insulin</td>
<td>13 years; 15 years</td>
</tr>
<tr>
<td></td>
<td>Cotinine</td>
<td>13 years; 14 years; 15 years</td>
</tr>
<tr>
<td></td>
<td>STIs and HIV</td>
<td>15 years</td>
</tr>
<tr>
<td></td>
<td>Vitamin D</td>
<td>9 years, 10 years, 11 years, 12 years, 13 years, 14 years; 15 years</td>
</tr>
<tr>
<td></td>
<td>Bone turnover markers</td>
<td>9 years, 10 years, 11 years, 12 years, 13 years; 14 years; 15 years</td>
</tr>
<tr>
<td>Child risk behaviour</td>
<td>Tobacco, drug and alcohol use; sexual experience; violence</td>
<td>10 years, 11/12 years, twice a year from 13-years</td>
</tr>
<tr>
<td>Sexually transmitted infections, including HIV</td>
<td>Through screening—reported sexual behaviour, pregnancy, ULE</td>
<td>Annually since 13 years</td>
</tr>
<tr>
<td>Genetics</td>
<td>DNA bank (n = 2200 child samples, 2200 one biological parent)</td>
<td></td>
</tr>
</tbody>
</table>
completeness and quality, entry, cleaning and the construction of analytical datasets.

Considerable energy goes into tracking and cohort maintenance. There are few street names and numbers in Soweto, informal housing springs up quickly and they confuse the visual landscape, many people have temporary work away from home and there are few fixed line telephones. Mobile phone usage is burgeoning, making it easier to contact caregivers and children. To date, attrition has been low in comparison to other developing country cohorts, and occurred largely in children’s early years. It is difficult maintaining better-off families who see little benefit in participating in a long-term study with twice yearly data collection, including fasting urine and blood draws. However, most African participants identify strongly with the study, and report being happy to participate if the information collected can be used to improve the quality of life of South Africans.

The study is unique and has been a source of reference for several major policy decisions in the country, particularly in the absence of other data. For example, information on children’s recognition of cigarette brands was used by the Minister of Health from 1997 to 1999 to push forward progressive tobacco legislation, preventing public advertising and the sale of cigarettes to minors. Early age of school entry revealed by longitudinal data was used by the Ministry of Education in 2004 to legislate a minimum age for school enrolment.

Looking backwards at what could have been done better is always with 20:20 vision. We could have over-enrolled White, Coloured and Indian families so that more systematic comparisons could have been made between the so-called Apartheid race groups. However, both at the time the study was initiated, and now, there is considerable ambivalence about treating these political classifications as biological, or even social, categories. Consultations with experts in longitudinal child studies advised us that the study would be maximally useful if it focused on the heterogeneity among a cohort of urban children, even if they were largely African. In the bone health study, an additional 120 White families were enrolled to ensure adequate power to test for ethnic differences in bone mass.

We certainly would benefit from having data that are more complete at each point. Circular migration and the difficulty in finding people means that some cases have missing data waves. Given the frequency of data collection, we can impute data from nearby time points for many variables, such as socioeconomic status or maternal education, but this is not possible for growth or developmental status. Since age 12, we have managed to achieve much more complete data at each round, and this has been helped by twice yearly data collection waves.

From the beginning, we have struggled to find the appropriate level of statistical expertise needed to do justice to longitudinal data. There are very few experienced applied statisticians in South Africa and statistics is regarded as a scarce skill throughout the science and technology sector. Collaborations with colleagues in the United States and the United Kingdom have brought with them the required expertise but, until very recently, this was ad hoc rather than an integral part of the staffing structure.

Can I get hold of the data? Where can I find out more?

The outline of the study, history and questionnaires are available on the website and collaborators can pull down data sets from the web through password access (http://www.wits.ac.za/birthto20). Collaborations are established through formal agreements with the principal investigators (richter@bsrc.ac.za), and include funds needed for the collaboration, staff requirements and mutual obligations, including the return of analytical datasets, notification of publications, etc. Our current grants cover costs involved in planned data collection and analysis. New initiatives that require additional staff or resources to collect or clean data have to be additionally funded.

Apart from a core group of South Africans at local universities and research institutions, such as Nigel Crowther from the University of the Witwatersrand who works on insulin resistance, we have active collaborations with Noel Cameron from Loughborough University on aspects of growth; Alan Stein and Paul Ramchadani from the University of Oxford on maternal depression, child adjustment and adolescent body image; Linda Clarke and Vicky Hosegood from the London School of Hygiene and Tropical Medicine on the role of fathers in children’s development; Marie-Louise Newell from the Institute of Child Health in London and the Africa Centre for Health and Population Studies in South Africa and Audrey Pettifor from the University of North Carolina, on adolescent sexual risk behaviour; the Lancet Child Development Series on a paper coordinated by Sally McGregor from the Institute for Child Health in London; and the Lancet Nutrition Series on a paper coordinated by Cesar Victora at the Universidade Federal de Pelotas.

Acknowledgements

From its inception, BT10 has been supported by the South African Medical Research Council. Other funders include the Institute for Behavioural Sciences at the University of South Africa. Since 1998, the Wellcome Trust has been BT20’s major funder, with additional support from the Human Sciences Research Council of South Africa, the Medical Research Council, the University of the Witwatersrand, the Mellon Foundation, the South-African Netherlands Programme on Alternative Development and the Anglo American Chairman’s Fund.

Special thanks are due to Sharon Fonn, Sylvia van Rensburg and Thea de Wet, who coordinated the study in its early phases. Many colleagues have been involved in various aspects of the study over the years, including Peter Cleaton-Jones, Peter Cooper, George Ellison, Dev Griesel, Allan Herman, Justice Hofmeyer, Carel Jusselmuiden, Angie Mathee, James McIntyre,

Dr Jean Golding, from the Avon Longitudinal Study of Pregnancy and Childhood (ALSPAC) Bristol University and Sir Michael Rutter at the University of London.
Nicky Padayachee, Krisela Steyn, Lucy Wagstaff and Dinky Levitt.
The study would never have survived or thrived without the dedication of staff, especially those field staff who have been with the study since it began: Barbara Moneypote, Eliza Tseou, Thabile Sibiya, Mantoia Langa and Andreas Phake.

References