

Equipping the community to measure children's height: the reliability of portable instruments

L D Voss, B J R Bailey

Abstract

Objective – To compare (1) the reliability of two expensive and two inexpensive measuring instruments, suitable for use in the community and (2) the reliability of experienced compared with inexperienced observers.

Design – (1) Ten children aged 5–12 years were each measured three times blindly, and in random order, by two experienced observers using four different portable instruments. (2) Four groups of four children aged 5–11 years were each measured three times blindly, and in random order by four experienced and one inexperienced measurer, using two different portable instruments.

Main outcome measures – The precision of height measurements made by different observers using different instruments, expressed in each case as the standard deviation of a single height measurement (SDsh_m).

Results – (1) No significant difference in precision was found between instruments, SDsh_m ranging from 0.22–0.34 cm. The two observers using apparently the same technique, did however record significantly different absolute heights. (2) No significant difference in precision was found between experienced and inexperienced observers.

Conclusion – Inexpensive height measuring equipment, once accurately installed, is no less reliable than the most expensive. Inexperienced observers can, with care, measure as reliably as those with long experience. Every effort should be made, however, to ensure that the progress of individual children is monitored not only by the same observer, but on a long term basis.

(*Arch Dis Child* 1994; 70: 469–471)

Screening for short stature at school entry has been shown to be of value in the identification of children who may have undiagnosed organic disease.¹ Repeated sequential height measurements that fall across the centile lines may also signal a need for investigation, whatever the stature of the child. The assessment of growth is a non-invasive means of monitoring the general health and wellbeing of every child in the community, but in the past has been performed poorly, leading to late diagnosis of serious growth disorders.²

Reliable height data must be both accurate and reproducible. We have previously reported the inaccuracy of many height

measuring instruments used in the community.³ This had little to do with cost and was mainly due to careless installation and a failure to calibrate. A spot check of 230 instruments revealed that one in seven was inaccurate by 1.0 cm or more, and that some were inaccurate by as much as 10 cm. Such errors are avoidable and easily checked. At the same time, we demonstrated the impossibility of obtaining perfectly reproducible measurements even when using a Harpenden stadiometer, the 'gold' standard used in many paediatric clinics.³

It is a common assumption that the most reliable data ought to be obtained by an experienced auxologist using the best instrument available. Indeed, the Hall report suggests that trained staff in specialist clinics, using expensive equipment, will produce the most reliable measurements.⁴ To test this hypothesis, we first compared the performance of four portable instruments, such as might be used in the community by health visitors or school nurses, and then we compared the performance of inexperienced with experienced measurers.

Subjects and methods

TRIAL 1: COMPARING INSTRUMENTS

Ten children, aged 5–12 years, heights 111.1 to 160.1 cm were each measured blindly three times by two experienced measurers (X and Y) on four accurately installed portable instruments. Two of the instruments were relatively inexpensive:

Leicester Height Measure (Child Growth Foundation, 2 Mayfield Avenue, Chiswick, London, price £34.70) – a light, plastic instrument, consisting of a footplate and four piece vertical ruler or scale and moveable headpiece. It needs no calibration, is quickly dismantled, packs flat and weighs 2.0 kg. (This instrument has been modified since the trial to make it a little more rigid.)

Minimetre (Child Growth Foundation, price £22.00) – a retractable, wall mounted metal tape, suspended either with plastic putty or, preferably, from a permanent hook. When fully extended, it should be self calibrating, but in practice, a metre rule is always required to make the final adjustment. A small plastic headpiece is pulled down onto the child's head and the whole instrument fits into a large pocket.

The two more expensive instruments were:

Free Standing Magnimetre (Raven Equipment Ltd, Castlemead Publications, 12 Little Mundells, Welwyn Garden City, Herts, price £320.00) – this rather bulky metal instrument

Wessex Growth Study,
Room CD53,
Southampton General
Hospital,
Southampton SO9 4XY
L D Voss

Faculty of
Mathematical Studies,
University of
Southampton
B J R Bailey

Correspondence to:
L D Voss.

Accepted 31 January 1994

has a folding tripod base, a fixed vertical scale requiring no calibration and removable magnetic sliding headpiece. The instrument is 136.0 cm in length when fully collapsed and weighs 8.5 kg.

Digirod (CMS Measuring Equipment, 18 Camden High Street, London, price £250.00) – this one piece, sturdy, self calibrating, all metal instrument, is telescopic in design and has a digital read out. It weighs 2.6 kg and is 75.0 cm long when fully retracted.

TRIAL 2: COMPARING MEASURERS

In a second trial, 16 children aged 5 to 11 years, heights 101.1 cm to 159.0 cm, were randomly divided into four independent groups (I–IV), each of which was allocated four specialist growth nurses, one inexperienced measurer, and two instruments. Within each group, the children were measured three times on each instrument by each measurer. The four groups were treated separately because the children were common to all observers within a group but not between groups. The two instruments used in each group were the Minimetre (as described earlier) and the Standard Magnimetre – similar to the portable model described earlier, but the scale is screwed permanently to the wall.

In both trials, the instruments were checked for accuracy using a metre rule both at the start and at completion. All measurements were made under standard experimental conditions, that is, blind, and in random order. Conventional anthropometric methods were used, except by the inexperienced measurers in the second trial, who were novices. They had been asked only to assist with the trial and had no idea that they would be required to make any measurements: no formal instructions were given although they were free to observe the technique of others, and some quite original styles were observed!

STATISTICAL METHODS

The Student's *t* test was used to compare means and Bartlett's test to compare standard deviations.

Results

(1) ACCURACY OF MEASUREMENTS

All instruments were accurate to within ± 0.1 cm. Repeated readings of a rigid 100 cm ruler ranged from 99.9 to 100.1 cm both at the start and end of the trials.

Table 1 Reproducibility of height measurements

Instrument	SDshm (cm)	95% Confidence interval for SD
Leicester Height Measure	0.28	0.21 to 0.41
Minimetre	0.22	0.17 to 0.33
Magnimetre	0.34	0.25 to 0.50
Digirod	0.33	0.25 to 0.49

The SDshm is based on measurements of 10 children. Each child was measured three times on each instrument. There was no significant difference in reproducibility of the two experienced observers on any instrument allowing their data to be pooled.

Table 2 Mean heights recorded by different measurers

Observer X	137.48 cm
Observer Y	137.19 cm
Difference	0.29 cm (SE=0.040 cm)
p Value	<0.001

Mean heights shown are for 10 children each measured three times on the four instruments by two observers.

(2) REPRODUCIBILITY OF MEASUREMENTS

Trial 1: comparing instruments

Table 1 shows the reproducibility of each instrument expressed as the standard deviation of a single height measurement (SDshm). There was no significant difference in precision among instruments, the SDshm ranging from 0.22 to 0.34 cm ($p=0.32$).

A significant difference was found, however, between the two experienced measurers in mean height obtained over all measurements. Observer X's mean height over 120 measurements was 0.29 cm greater than that of observer Y (table 2). Moreover, a highly significant interaction between measurers and children was observed ($p<0.01$), due almost entirely to measurer X obtaining greater heights than Y for nine of the children but not the 10th, who was in fact the tallest child and may have been difficult to stretch.

Trial 2: comparing measurers

Table 3 shows for each group in turn, the SDshm obtained by each observer. The results from the two instruments have been pooled in every case, as there were no more differences between instruments in their standard deviations than might be expected by chance. Within each group there were barely any differences in precision among the five measurers (four experienced, one inexperienced); *p* values=0.49, 0.044, 0.17, and 0.86.

It was again noted, however, that, even among experienced measurers, different observers obtained significantly different mean heights for the same group of children.

Discussion

We have shown previously that the inaccurate installation of equipment is a frequent source of error in height measurement, but this can be avoided, as in these trials, by careful calibration using a standard rule before and after every measurement session.³ Instruments that require no calibration or are self calibrating are

Table 3 Trial 2: comparing measurers (SDshm in cm)

Group I		Group II	
Experienced observers	0.44	Experienced observers	0.18
	0.34		0.26
	0.32		0.24
	0.39		0.21
Inexperienced observer	0.24	Inexperienced observer	0.27
Group III		Group IV	
Experienced observers	0.34	Experienced observers	0.38
	0.47		0.30
	0.29		0.32
	0.26		0.37
Inexperienced observer	0.49	Inexperienced observer	0.37

The SDshm for each observer in each group is based on measurements of four children each measured three times on two different instruments: a Magnimetre and a Minimetre.

clearly easier to use and more reliable than the Minimetre which has to be recalibrated each time it is moved. Unless permanently positioned it is not the ideal instrument for monitoring the growth of individual children.

On the other hand, little can be done to minimise poor reproducibility. Measurement of animate subjects will always have a degree of imprecision. The SD for a single height measurement was found in this study to vary little between instruments or observers, confirming previous findings by ourselves and others.^{3,5} This should not be surprising as we have already shown the variability in height measurement to be due largely to the flexibility or ever changing posture of the child. Less than 10% of the variance is attributable to the measurer or instrument.³ In terms of precision, therefore, the least expensive instruments, that is, the Minimetre and the Leicester Height Measure, perform as well as the more expensive models, the free standing Magnimetre and the Digirod. All instruments compare favourably with the Harpenden stadiometer tested earlier.³

In view of the variance described above it is also not surprising that the reproducibility or precision of the *inexperienced* measurers is able to match that of the experienced auxologists. That is not to say that training and experience are not to be recommended. Where children are remeasured at a later date, an experienced auxologist will use the same well rehearsed technique, whereas an inexperienced measurer may well employ a new technique with misleading results.

Measurement techniques clearly differ even between experienced observers, each one stretching or positioning the child in a different way. Although this may not affect the *precision* or reproducibility of their measurements, it can result in the recording of significantly different *mean* heights when different measurers measure the same children. Where the same health worker always monitors the same child, there is no difficulty. Misleading data could arise if a child were to be measured on two occasions by two different people. As much as 1 cm of any apparent increment, a large proportion of annual growth, could be attributable simply to the difference between two observers.

The notion that each child has a true or absolute height which could in theory be ascertained by an expert using the best equipment available, is clearly a myth. Each child has only an average height at any one time. Given a typical SD of between 0.2 to 0.4 cm as found in these and previous trials, it is only possible, *at best*, to measure height to within approximately ± 0.4 cm. There are clearly serious implications for the interpretation of short term growth data⁶ on which, unfortunately, the

diagnosis and management of growth problems is increasingly based.⁷⁻¹⁰ Height data can be reliable, but only if regular measurements are begun at an early age,¹¹ and made over a sufficient period of time in order to establish a trend.⁶ There are no gold standards and no short cuts in the assessment of growth.

It is reassuring, however, that the monitoring of children's heights has been shown to be a viable proposition at community level. One careful, consistent observer, using an inexpensive, easily calibrated instrument, can collect data no less reliable than an experienced auxologist operating the most expensive equipment.

CONCLUSIONS AND RECOMMENDATIONS

(1) Reliable growth data requires neither expensive equipment nor lengthy training.

(2) Most of the variance in height measurement is unavoidable and due to the changing posture and flexibility of the subject. Perfect precision is an impossible goal but consistency of technique should optimise the reproducibility of the measurements.

(3) Errors due to inaccuracy may be considerable, but are avoidable if a standard rule is always used to check calibration.

(4) Interobserver differences in technique can be a major source of error, and must be borne in mind where a different person has taken over the monitoring of a child.

(5) Long term monitoring, begun at an early age, is now both feasible and affordable and should be standard practice in any child health surveillance programme.

We are grateful to Dr P Swift (who designed the Leicester instrument) and Dr P Betts for their helpful comments. We are also grateful to Dr E S McCaughey, Jean Mulligan and the many children and specialist growth nurses who took part in the trials.

LDV is generously supported by a grant to the Wessex Medical Trust from Kabi Pharmacia UK and Sweden AB.

- Voss LD, Mulligan J, Wilkin TJ, Betts PR. Poor growth in school entrants as an index of organic disease: the Wessex Growth Study. *BMJ* 1992; **305**: 1400-2.
- Aynsley-Green A, Macfarlane JA. Method for the earlier recognition of abnormal stature. *Arch Dis Child* 1983; **58**: 535-7.
- Voss LD, Bailey BJR, Cumming K, Wilkin TJ, Betts PR. The reliability of height measurement (the Wessex Growth Study). *Arch Dis Child* 1990; **65**: 1340-4.
- Hall DMB, ed. *Health for all children*. 2nd Ed. Oxford: Oxford Medical Publications, 1991.
- Ahmed ML, Yudkin PL, Macfarlane JA, McPherson K, Dunger DB. Are measurements of height made by health visitors sufficiently accurate for routine screening of growth? *Arch Dis Child* 1990; **65**: 1345-8.
- Voss LD, Wilkin TJ, Bailey BJR, Betts PR. The reliability of height and height velocity in the assessment of growth (the Wessex Growth Study). *Arch Dis Child* 1991; **66**: 833-7.
- Brook CGD. Treatment of growth deficiency. *Clin Endocrinol (Oxf)* 1988; **34**: 65-72.
- Hindmarsh PG, Brook CDG. Auxological and biochemical assessment of short stature. *Acta Paediatr Scand [Suppl]* 1988; **343**: 73-6.
- Milner RDG. Which children should have growth hormone therapy? *Lancet* 1986; **i**: 483-5.
- Ahmed ML, Allen AD, Sharma A, Macfarlane JA, Dunger DB. Evaluation of a district growth screening programme: the Oxford growth study. *Arch Dis Child* 1993; **69**: 361-5.
- Betts PR, Voss LD, Bailey BJR. Measuring the heights of very young children. *BMJ* 1992; **304**: 1351-2.