



TwinLife

TwinLife Technical Report Series

No. 06, December 2017

# Documentation TwinLife Data: Height, Weight, and BMI

by Franziska Lenau & Elisabeth Hahn

Franziska.Lenau@uni-saarland.de





# TwinLife

**Franziska Lenau & Elisabeth Hahn**

**Documentation TwinLife Data: Height, Weight, and BMI**

TwinLife Technical Report Series No. 06

Project TwinLife “Genetic and social causes of life chances”

Bielefeld, December 2017

TwinLife Technical Report Series

General Editors: Martin Diewald, Rainer Riemann and Frank M. Spinath

ISSN 2512-403X

This publication has been funded by the German Research Foundation (DFG).

TwinLife Technical Reports are refereed scholarly papers. Submissions are reviewed by the general editors before a final decision on publication is made.

The Technical Report Series is a forum for presenting technical works (e.g. data documentation, field reports etc.) in progress. Readers should communicate comments on the manuscript directly to the author(s).

The papers can be downloaded from the project website:

<http://www.twin-life.de/en/twinlife-technical-report-series>

TwinLife “Genetic and social causes of life chances”

University of Bielefeld

Faculty of Sociology

PO Box 100131

D-33501 Bielefeld

Germany

Phone: +49 (0)521 106-4309

Email: [martin.diewald@uni-bielefeld.de](mailto:martin.diewald@uni-bielefeld.de)

Web: <http://www.twin-life.de/en>



<b>Content</b>	<b>Page</b>
1. Description of measurement and general procedure	2
2. Corrections of multiple and extreme values	2
· Multiple values	2
· Implausible and extreme values	3
3. Height, weight, and BMI variables in the <i>TwinLife</i> dataset	3
<b>References</b>	4
<b>Appendix</b>	
TABLE 2. <i>Multiple values with differences of 10 or greater.</i>	5
TABLE 3. <i>Implausible values.</i>	7
TABLE 4. <i>Extreme values.</i>	8

## 1. Description of measurement and general procedure

In the *TwinLife* study, current height [in cm] and weight [in kg] were assessed in an open-ended format<sup>1</sup>. For children aged up to 13 years one parent was asked to indicate the offspring's current height and weight. For parents as well as children and adolescents aged 14 years or older, current height and weight were measured via self-report. Within the household interview, participants were given the opportunity to measure or weight themselves, if necessary.

Information on height and weight should be given only by one informant, i.e. either via self-report or one parent-report. This occurs in 16,109 of 16,284 cases for height (98.9 %) and 15,827 of 16,997 cases for weight (98.9 %). For the remaining 175 cases (1.1 %), respectively 170 cases (1.1 %), questions were answered at least twice (i. e., via self- and parent-report or via two parent-reports). Multiple values occurred only in the offspring's data, for children aged 3 to 21 years ( $M_{age} = 11.9$ ). Multiple height values were given for  $n = 34$  twins and  $n = 141$  siblings. Regarding weight, multiple answers were given for  $n = 37$  twins and  $n = 133$  siblings.

## 2. Corrections of multiple and extreme values

### Multiple values

Two “count variables” were generated, indicating the number of height (*nht*) or weight (*nwt*) values for one person in the data set. In the presence of multiple values, absolute differences below 10 centimeters or kilogram, multiple height or weight values were averaged to create final height (*hgt*) or weight (*wgt*) variables.

In cases, in which height ( $n = 174$ ) or weight ( $n = 169$ ) values were available twice, absolute differences between both height (*dht*) or weight (*dwt*) values were calculated. For one case (*pid* = 224233200) it occurred that the self-report (157 cm, 44 kg) as well as the maternal-report (158 cm, 47 kg) and also the step-fathers report (156 cm, 44 kg) were filled out. In this case, the mean difference was calculated (*dht* = 1.3 cm; *wht* = 2.0 kg).

However, when differences of 10 centimeter or greater ( $n = 9$ ) respectively 10 kilogram or greater ( $n = 8$ ) occurred, cases were double checked: Both values were averaged to create final height or weight variables, if both were plausible with regard to the person's age (compared to the population) as well as to height and / or weight of the family members. If only one value occurred to be realistic, the other was not taken into account. Cases concerned are documented in the appendix (TABLE 2).

---

<sup>1</sup> Despite current height and weight, we also assessed the offspring's height and weight at nine medical check-ups from immediately after birth to school-age in the *TwinLife* study.

## Implausible and extreme values

In  $n = 13$  cases, height and / or weight were classified as *implausible values* (see appendix, TABLE 3). In three cases, values were declared as missing values in the final height (*hgt*) or weight (*wgt*) variables. In two cases, height and weight values were interchanged in the final variables. In eight cases, values seemed to be incorrect because of typing errors, which were corrected in the final height (*hgt*) or weight (*wgt*) variables. Some cases ( $n = 9$ ) were classified as plausible *extreme values* (see appendix, TABLE 4), after checking each case and, if necessary, double-checking height and / or weight of the family members. For these cases, values were maintained in the final height (*hgt*) or weight (*wgt*) variables.

### 3. Height, weight, and BMI variables in the *TwinLife* dataset

In conclusion to the described procedure, different height, and weight variables can be found in the *TwinLife* dataset. First, there are several *original, i.e. uncorrected variables* for height and weight from self-reports (*bdy0100*, *bdy0200*) and parent-reports (*bdy0100t*, *bdy0100u*, *bdy0100s*). Suffixes were used to indicate whether the information in the parent-report was given by mother (*\_m*), father (*\_f*), stepmother (*\_n*), or stepfather (*\_g*). Furthermore, there are additional *generated variables* indicating whether there were multiple height (*nht*) or weight (*nwt*) values and, if so, specifying the difference between these multiple values (*dht*, *dwt*). Finally, there are *corrected variables* for height (*hgt*), and weight (*wgt*), which should be used as the “final variables”.

Based on the entries for height and weight, *Body Mass Index* (BMI; Warschburger, 2017) was calculated as follows:  $BMI = \text{Weight [kg]} / \text{Height}^2 [\text{m}^2]$ .

TABLE 1: Overview about height, weight and BMI variables in the *TwinLife* dataset

Type of variable	Variable name	Description
Original variables, uncorrected	<i>bdy0100</i>	= Self-report: body height in cm
	<i>bdy0100t_m/f/n/g</i>	= Parent-report: height of twin 1 in cm
	<i>bdy0100u_m/f/n/g</i>	= Parent-report: height of twin 2 in cm
	<i>bdy0100s_m/f/n/g</i>	= Parent-report: height of sibling in cm
	<i>bdy0200</i>	= Self-report: weight in kg
	<i>bdy0200t_m/f/n/g</i>	= Parent-report: weight of twin 1 in kg
	<i>bdy0200u_m/f/n/g</i>	= Parent-report: weight of twin 2 in kg
Generated auxiliary variables	<i>bdy0200s_m/f/n/g</i>	= Parent-report: weight of sibling in kg
	<i>nht</i>	= Number of height values given for one person
	<i>nwt</i>	= Number of weight values given for one person
	<i>dht</i>	= Difference between several height values
Generated variables, corrected	<i>dwt</i>	= Difference between several weight values
	<b><i>hgt</i></b>	= <b>Final height variable: corrected</b>
	<b><i>wgt</i></b>	= <b>Final weight variable: corrected</b>
	<b>BMI</b>	= <b>Final BMI variable: corrected</b>

## References

Warschburger, P. (2017). Body-Mass-Index (BMI). In M. A. Wirtz (Hrsg.), *Dorsch – Lexikon der Psychologie*. Retrieved from, <https://portal.hogrefe.com/dorsch/body-mass-index-bmi/>

Glenday, C. (Ed.). (2014). *Guinness world records 2014*. Bantam.

## Appendix

TABLE 2. Multiple values with differences of 10 or greater.

Variable concerned ▸ Person concerned	pid (Person ID)	Informant	Height [in cm]	Weight [in kg]	BMI_ original	Handling	BMI_final
Height & Weight ▸ 6-year-old sibling	239593200	Self Mother	175.0 119.0	70.0 20.0	22.86 14.12	<i>Self-reported height &amp; weight implausible for 6-year-old child: → Maternal-reported values</i>	14.12
Height & Weight ▸ 10-year-old sibling	445592200	Self Mother	165.0 146.0	55.0 34.0	20.20 15.95	<i>Both values possible → Averaged value: Height = 155.5, Weight = 44.5</i>	18.40
Height ▸ 13-year-old sibling	333657200	Self Mother	140.0 160.0	40.0 45.0	20.41 17.58	<i>Both values possible → Averaged value: Height = 150.0, Weight = 42.5</i>	18.89
Height ▸ 13-year-old sibling	345125200	Self Mother	180.0 163.0	61.0 60.0	18.83 22.58	<i>Check – Height &amp; weight of family members ▸ Twin 1 (345125001): Height = 199, Weight = 70 ▸ Twin 2 (345125002): Height = 198, Weight = 76 ▸ Mother (345125300): Height = 165, Weight = 70 Both values possible → Averaged value: Height = 171.5, Weight = 60.5</i>	20.57
Height ▸ 13-year-old sibling	262672200	Self Mother	148.0 159.0	42.0 42.0	19.17 16.61	<i>Both values possible → Averaged value: Height = 153.5, Weight = 42.0</i>	17.83
Height ▸ 13-year-old sibling	230277200	Self Mother	160.0 150.0	37.0 45.0	14.45 20.00	<i>Both values possible → Averaged value: Height = 155.0, Weight = 41.0</i>	17.07
Height ▸ 13-year-old sibling	177834200	Self Mother	150.0 140.0	33.0 35.0	14.45 17.86	<i>Both values possible → Averaged value: Height = 145.0, Weight = 34.0</i>	16.17
Height ▸ 11-year-old sibling	274495200	Father Mother	132.0 148.0	33.0 38.0	18.94 17.35	<i>Both values possible → Averaged value: Height = 140.0, Weight = 35.5</i>	18.11
Height ▸ 8-year-old sibling	253910200	Father Mother	140.0 150.0	35.0 35.0	17.86 15.56	<i>Both values possible → Averaged value: Height = 145.0, Weight = 35.0</i>	16.65
Weight ▸ 13-year-old sibling	210076200	Self Mother	175.0 170.0	75.0 60.0	24.49 20.76	<i>Both values possible → Averaged value: Height = 172.5, Weight = 67.5</i>	22.68

<b>Variable concerned</b> ‣ Person concerned	<b>pid</b> (Person ID)	<b>Informant</b>	<b>Height</b> [in cm]	<b>Weight</b> [in kg]	<b>BMI_</b> original	<b>Handling</b>	<b>BMI_final</b>
Weight ‣ 13-year-old sibling	221865200	Self Mother	165.0 169.0	50.0 65.0	18.37 22.76	<i>Both values possible → Averaged value:</i> Height = 167.0, Weight = 57.5	20.62
Weight ‣ 13-year-old sibling	286672200	Self Mother	159.0 164.0	38.0 49.0	15.03 18.22	<i>Both values possible → Averaged value:</i> Height = 161.5, Weight = 43.5	16.68
Weight ‣ 13-year-old sibling	294896200	Self Mother	182.0 178.0	60.0 50.0	18.11 15.78	<i>Both values possible → Averaged value:</i> Height = 180.0, Weight = 55.0	16.98
Weight ‣ 13-year-old sibling	255429200	Self Mother	175.0 170.0	65.0 75.0	21.22 25.95	<i>Both values possible → Averaged value:</i> Height = 172.5, Weight = 70.0	23.52
Weight ‣ 13-year-old sibling	296404200	Self Mother	158.0 157.0	75.0 85.0	30.04 34.38	<i>Both values possible → Averaged value:</i> Height = 157.5, Weight = 80.0	32.25



TABLE 3. *Implausible values.*

<b>Variable concerned</b> ‣ Person concerned	<b>pid</b> (Person ID)	<b>Informant</b>	<b>Height</b> [in cm]	<b>Weight</b> [in kg]	<b>BMI_</b> original	<b>Handling</b>	<b>BMI_</b> final
Height & Weigh ‣ Twin (cohort 4)	433249002	Self	1.0	1.0	10000.00	<i>No correction: Declared as missing values</i>	Missing
Height ‣ 6-year-old sibling	138527001	Father	175.0	Missing	Missing	<i>No correction: Declared as missing value</i>	Missing
Weight ‣ 12-year-old sibling	272800002	Mother	Missing	214.0	Missing	<i>No correction: Declared as missing value</i>	Missing
Height & Weight ‣ 21-year-old sibling	434771200	Self	42.0	150.0	850.34	<i>Correction: Height &amp; weight interchanged</i>	18.67
Height & Weight ‣ Mother	229005300	Self	74.0	171.0	312.27	<i>Correction: Height &amp; weight interchanged</i>	25.31
Height ‣ Mother	411477300	Self	54.0	48.0	164.61	<i>Correction: Height = 154.0</i>	20.24
Weight ‣ Twin (cohort 1)	114097001	Father	110.0	188.0	155.37	<i>Check – Twin (114097002): Height = 114, Weight = 20 → Correction for 114097001: Weight = 18.8</i>	15.54
Weight ‣ Twin (cohort 1)	126339001	Mother	105.0	176.0	159.64	<i>Check – Twin (126339002): Weight = 18, Height = 108 → Correction for 126339001: Weight = 17.6</i>	15.96
Weight ‣ Twin (cohort 1)	164528001	Mother	110.0	191.0	157.85	<i>Check – Twin (164528002): Height = 119, Weight = 27 → Correction for 164528001: Weight = 19.1</i>	15.79
Weight ‣ Twin (cohort 1)	115568002	Mother	106.5	145.0	127.84	<i>Check – Twin (115568001): Height = 106, Weight = 14 → Correction for 115568002: Weight = 14.5</i>	12.78
Weight ‣ Twin (cohort 1)	173154001	Mother	107.0	107.0	93.46	<i>Check – Twin (173154002): Height = 112, Weight = 20 → Correction for 173154001: Weight = 17.0</i>	14.85
Weight ‣ Twin (cohort 1)	115584002	Father	108.0	107.0	91.74	<i>Check – Twin (115584001): Height = 110, Weight = 14 → Correction for 115584002: Weight = 17.0</i>	14.57
Weight ‣ Twin (cohort 2)	299846002	Mother	130.0	128.0	75.74	<i>Check – Twin (299846001): Height = 130; Weight = 28 → Correction for 299846002: Weight = 28.0</i>	16.57

TABLE 4. *Extreme values.*

<b>Variable concerned</b> ‣ Person concerned	<b>pid</b> (Person ID)	<b>Informant</b>	<b>Height</b> [in cm]	<b>Weight</b> [in kg]	<b>BMI_</b> original	<b>Handling</b>	<b>BMI_</b> final
Height ‣ Twin (cohort 4)	463889002	Self	115.0	Missing	Missing	<i>Check – Height &amp; weight of family members</i> ‣ <i>Twin (463889001): Height = 158, Weight = 53</i> ‣ <i>Mother (463889300): Height = 156, Weight = 82</i> → <i>No correction: possible extreme value</i>	Missing
Height & Weight ‣ Twin pair (cohort 2)	233608001 233608002	Mother Mother	176.0 176.0	106.0 104.0	34.22 33.57	<i>Check – Height &amp; weight of family members</i> ‣ <i>Mother (233608300): Height = 180, Weight = 129</i> ‣ <i>13-year-old sibling (233608200): Height = 184, Weight = 126</i> → <i>No correction: values possible</i>	34.22 33.57
Height ‣ Mother	451555300	Self	126.0	61.0	38.42	<i>Both values possible → No correction</i> ‣ <i>The shortest woman living measures 62 cm [Guinness World Records, 2014]</i>	38.42
Weight ‣ Twin (cohort 3)	388532001	Self	100.0	64.0	64.00	<i>Check – Twin (388532002): Height = 155, Weight = 70</i> → <i>No correction: possible extreme value</i>	64.00
Weight ‣ Twin (cohort 3)	312512001	Self	167.0	176.0	63.11	<i>Check – Twin (312512002): Height = 168, Weight = 58</i> → <i>No correction: Possible extreme value</i>	63.11
Weight ‣ Twin (cohort 4)	440194002	Self	155.0	165.0	68.68	→ <i>No correction: possible extreme value</i>	68.68
Weight ‣ Twin (cohort 4)	445379002	Self	162.0	163.0	62.11	→ <i>No correction: possible extreme value</i>	62.11
Weight ‣ Twin (cohort 4)	467878001	Self	165.0	140.0	51.42	→ <i>No correction: possible extreme value</i>	51.42