1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

**Establishing the variation of blood metrics within UK Biobank
to inform cell therapy manufacturing processes**

J Thurman-Newell, J Petzing, D J Williams

Centre for Biological Engineering, Loughborough University,
Loughborough, Leicestershire, LE11 2QZ, UK

Corresponding author: J Petzing

j.petzing@lboro.ac.uk

Loughborough University

25 **Abstract**

26 **Aim:** Understanding blood component variation as a function of healthy population
27 metrics is necessary to inform biomanufacturing process design. **Methods:** UK
28 Biobank metrics were examined for variation in White Blood Cell (WBC) count as an
29 analogue to potential manufacturing starting material input. **Results:** WBC variation
30 of four orders of magnitude (6.65×10^9 cells / litre) was found. Variation increased with
31 age, increased with weight up to 80 kg then decreased. Health state showed a
32 greater absolute number of participants with elevated count. Female range was
33 greater than male. Cell count / distributions were different between centres.
34 **Conclusions:** This variation and range of process input signals a requirement for
35 new strategies for manufacturing process design and control.

36

37 **Keywords:**

38 UK Biobank, White Blood Cell count, variation, manufacturing, control

39

40

41 5,900 words in core text, excluding Tables, References and Abstract

42 115 words for Abstract

1.0 Introduction

Advanced therapy medicinal products (ATMPs) such as cell and tissue therapies have changed the potential of the treatment landscape over the last twenty years, from a patient perspective and a clinical perspective but more recently from a potential manufacturing perspective. Autologous and allogeneic regenerative medicine therapeutics offer powerful and personalised treatment options that are typically delivered on an individual basis in a clinical setting. Inevitably the demand for these types of therapeutics, be they chronic or acute, life-changing or life-saving, far outstrips the capability for current production. Consequently, there is a growing need for alternative methods of production, such as centralised or distributed manufacture, by clinics, by small to large companies, and a delicate balance by Regulatory authorities [1,2] between therapy risk and benefit return for patients.

Recent releases from the US Food & Drug Administration (FDA) and the European Medicines Agency (EMA) further recognise the complexity of manufacturing cell-based therapies. The 2017 GMP Guidelines for ATMPs [3], for example, confirm these therapeutics are at the forefront of innovation and this unexplored territory impacts the biomanufacturing process. New biomanufacturing models are emerging to address the specific challenges of ATMPs, including decentralised manufacturing for autologous products at multiple sites in hospital settings. As the EMA Guidelines observe: “*ATMPs manufactured under a decentralised system as manufacturing in multiple sites increase the risk of variability for the product*” – however, for the products of the future, especially for common chronic or general illness, multiple site manufacture may be the only viable solution.

The potential for multi-site biomanufacturing has been further discussed in a recent statement by the FDA [4]. This introduces a route to licence for “*groups of investigators or small firms that are able to consistently follow a common manufacturing and clinical protocol but that may not have access to the patient populations or infrastructure needed to conduct separate development programs*”. Products from the group would be co-developed, but licensed for individual sites

74 which would by implication be different clinical centres – this brings into question the
75 established traditional tenet of multi-site comparability.

76
77 Large-scale production of cell therapies, whether by single or multiple sites,
78 introduce challenges to be solved to achieve the delivery of safe, viable therapeutics.
79 These challenges include the translation of the process from T-flask to bioreactor,
80 inline measurement and assessment of the cell journey, concerns around infections
81 and asepsis, correct cell differentiation, supply logistics, business considerations,
82 and the quality and variation of starting materials [5–7]. Whilst it is recognised that
83 variation may be caused by the biomanufacturing process, little public domain
84 information is available to quantify these contributions. Instead, it is the latter
85 element (variation of starting materials) that has formed the focus of recent attention
86 and our reporting here.

87
88 Manufacturing processes and control paradigms have been traditionally predicated
89 on a product specification, with minimal variation of the product and the capability to
90 control the manufacturing process using product related limits. Automation within
91 traditional manufacture that has adhered to this philosophy has led to in-process,
92 statistical process control. Most typically, process control approaches accommodate
93 a few percent variation in process metrics using software control or manual quality
94 control methods. Whilst there is a requirement to translate, scale-up and scale-out
95 the production of cell therapies, the biomanufacturing processes need to understand,
96 and either accommodate or control the issues associated with, and context of, using
97 cells as products. To achieve this will require collaboration and understanding
98 between the biological and manufacturing disciplines and the interface between
99 them.

100
101 Furthermore, it is important to understand and manage the variation of the
102 biomanufacturing process, both at the centre of the statistical distribution, and at its
103 extremes (i.e. outliers), in addition to understanding if variation is predicated on
104 normal probability distributions, or skewed non-parametric distributions. Informal

discussions in the UK have highlighted that the Regulator would want to focus on the behaviour of any biomanufacturing process at its extremes and recognise that any extreme outlier would be a patient requiring treatment and thus could not be disregarded. This is a critical consideration in the design of the cell therapy biomanufacturing process and can only be addressed with a clearer understanding of the magnitude and distribution of the extremes that could be encountered in any given population.

Recent research has started to examine the key issues at this interface, specifically the potential for variation within an Haematopoietic Stem Cells Therapy (HSCT) exemplar. A meta-analysis of literature across a 30-year window identified that variation of prescribed cell dose could be up to 6 orders of magnitude [8-10], depending on clinical centre, country, patient and medical team. Contextually, conventional manufacturing process control would normally be defined as 10% to 30% of the defined process limits (industry and product specific) i.e. typically less than 0.1 orders of magnitude. To date, traditional manufacturing process control processes typically cannot accommodate process variation in the orders of magnitude, and if these control methods are to be used for biomanufacturing processes, then potentially they will not be able to determine correct process parameters with respect to the cell therapy specification. This statement is also made in the context that as to date, there is very limited consensus about in theatre dose specification.

Whilst this research is insightful, it has been derived from a retrospective analysis. A more refined definition of variation within the starting material for HSCT was recently investigated from a single-site clinical centre [10,11]. It was hypothesised that a single, exemplary clinical centre would apply better control to sourcing and processing of starting material, and thus provide a more contemporary definition of current processing control limits. Through investigation of approximately 400 patient records across a one-year period of treatments at a specialist cancer treatment centre in the USA, it was identified that the variation at this single centre was reduced and better controlled – however, the variation was still found to be up to 3

orders of magnitude, with similarities between predominantly autologous product metrics. This variation was seen to be a function of initial collection (peripheral blood or bone marrow aspiration), mobilisation regimes (when used), pooling methods (when used), and whole blood processing methods.

Whilst gaining comprehension from a larger single site patient sample size would be desirable, it was concluded from this analysis of approximately 400 patient records [10, 11], that given the international reputation of the clinical centre the variation encountered would likely be representative of a common clinical centre. Although this informs us of the metrics associated with predominantly autologous blood and apheresis cell collection of unwell patients, it does not provide a perspective of the wider, relatively healthy population and the potential to benchmark the general population.

The aim of this research was for the first time to quantify this variation, encountered in a general, healthy population, as characterised by the UK Biobank [12]. This represents an opportunity to study the baseline population in the UK with respect to blood measurands, in particular White Blood Cell (WBC) counts for overall biological variation and potential influencing factors. An elevated WBC can be seen as a risk factor for various diseases (e.g. Cardiovascular, Coronary, Ischemic Stroke) [13] and it is necessary to understand the physiological relationship between patient or donor health and cell quality or efficacy, so that suitable donors for cell therapies can be identified and procured, providing more control over variation in the blood product at the biomanufacturing stage.

UK Biobank also presented an opportunity to investigate specific patient/donor metrics that were previously unavailable within both the meta-analysis [8,9] and the single clinical centre analysis [11], such as medical history, lifestyle specifics such as smoking, and consistent physical measurements such as weight and exercise. This permitted exploration of whether the existing data has sufficient explanatory power to

167 assist in the stratification of likely patients or patient groups in order to manage the
168 consequences of variation.

169

170 The authors consider in this work that WBC count is an appropriate illustrative
171 analogue for the demonstrating the potential variation (and the ranges of variation) in
172 starting materials for cell therapies, given that at the time of analysis the UK Biobank
173 dataset did not provide any more detailed quantification of cell function/identity (e.g.
174 CD markers). However, in this context and within the analysis reported here, it is not
175 possible (and consequently not the purpose) to directly correlate potential variation
176 of WBC within the UK Biobank participants with the previously defined CD34+
177 variation found in sick patient populations [8 – 11], other clinical studies, or existing
178 cell therapy manufacturing examples. The key purpose of this work is to demonstrate
179 to the cell therapy biomanufacturing community the possibility of blood metric
180 variation within a large scale nominally healthy population, that may beneficially
181 influence (primarily) allogeneic starting material selection for a range of potential
182 therapies.

183

2.0 Methodology

2.1 UK Biobank metrics

UK Biobank is a unique health resource and a UK charity [12], established with the aim of improving the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses. It consists of a databank and a biological sample bank from over 500,000 recruited individuals, between the ages of 40 and 69, with donors sampled between 2006 and 2010, and is intended to be a resource for investigating potential causes of disease development. The UK Biobank project collected biological samples and questionnaire data from volunteers in 22 centres around the UK [14]. The qualitative data and the biological samples were then transferred to a central processing and storage facility in Manchester (UK), and represents one of the very few international (notionally) healthy population biological resources.

		Participants (n)	Minimum	Maximum	Mean	Median
Unique ID		502,664	-	-	-	-
Gender	Male	229,182	-	-	-	-
	Female	273,467	-	-	-	-
Birth Year		502,649	1934	1971	1952	1950
Age		502,649	45	82	64.46	66.00
Weight (kg)		499,874	30.0	197.7	78.1	76.4
BMI		499,543	12.12	74.68	27.43	26.74

Table 1: Summary of Participants within UK Biobank [12]

A summary of the contents of the UK Biobank database is shown in Table 1. The resolution of the UK Biobank ($n = 502,664$) allows for analysis of participant / patient specific characteristics such as lifestyle, medical history and other physical metrics (e.g. weight), as well as the investigation of possible influencing factors, such as ethnicity or diet, which may prove to be risk factor when it comes to particular diseases and conditions. Blood specific measurements such as WBC counts were recorded at three time intervals: initial testing, repeat testing and imaging, these being three successive phases of blood sample analysis completed by UK Biobank [15]. Analysis

207 within the work reported here is based upon the counts recorded at initial testing,
208 because this time period contains the largest proportion of the dataset.

209

210 The application for this work focused specifically on the physical measurements of
211 the donor, the questionnaire results – particularly those focused on diet and health –
212 and the results of the whole blood test carried out on corresponding sample. Figures
213 1 to 4 represent those specific measurements from the UK Biobank database
214 required for this analysis. These results were chosen based on the experiences with
215 the medical literature and subsequent meta-analysis [8,9], and single site clinical
216 process records [11]. These results included previous medication (Figure 1) physical
217 metrics such as diet, weight and age (Figure 2), sleep duration, ECG and pulse rates
218 (Figure 3).

219

220 At the point that this investigation was carried out, UK Biobank did not directly
221 measure cell markers that specifically identified stem cells such as Haematopoietic
222 Stem Cells (HSCs) [15]. CD34+ cell counts would have been appropriate to directly
223 compare the relatively healthy population of the UK Biobank to the relatively unwell
224 populations in multi and single centre clinical analysis. UK Biobank does however
225 record red blood cell counts, WBC counts and the specific subsets of WBCs such as
226 leukocytes and monocytes although the latter two components were not required to
227 be featured in our analysis (Figure 4). As previously defined variation was therefore
228 explored as a function of WBCs as a surrogate to HSCs.

Figure 1: Participant Metrics – General, Ethnicity and Medical Health

Figure 2: Participant Metrics – Physical Body Measurements

Figure 3: Participant Metrics – Diet and Exercise

Figure 4: Blood Composition Metrics

The application to the UK Biobank consisted of two types of metrics that were requested; participant metrics and blood composition metrics;

- Participant Metrics: These measurements were either physical, or as part of the questionnaire, and were requested to determine whether these had any effect on the variation measured in the corresponding blood measurements.
- Blood Composition Metrics: These were characteristics specifically of the tested donor's blood content including white and red blood cell counts. These were the measurands being directly examined to quantify variation and may have been affected by the physical and lifestyle metrics of the previous category.

Pregnancy and menopause were an inherent integral of this analysis because of the biological changes that come as a result of these two life phases, and whether the advent of these phases has an effect on the variation encountered as a result – this may have implications for future cellular therapy, such as using female donors over, or under, a certain age for allogeneic therapy. However, finer resolution of analysis to specifically investigate participants identifying pregnancy and menopause was not required at this stage.

Assessment centre, and attending date were included to examine the effect of geographical location on the variation and spread of biological data; i.e. was there more biological variation on the coast compared to in-land? If certain conditions were indicative of more or less variation or cell quality, were there greater or lesser incidences of those conditions depending on location? What effect does living in a rural area have, if any? More likely, these assessment centres represent sub-sets of the population, with particular ethnic and genetic makeups – much like the reasoning behind setting up blood donation centres in specific locations to capture more of a specific, rarer blood group or type.

Job code, and other similar metrics shown in the top right corner of Figure 1, were included with the intention of determining whether sedentary jobs such as office work affect cell characteristics differently to more manual or active jobs.

Medication, treatment, operation code and other metrics shown in the bottom right corner of Figure 1 were included because of previous experience as to the effect of the mobilisation and conditioning regime. These were considered to have been likely to be the most influential factors for amount and quality of cells, and were therefore included to examine their subsequent effect, if any, on biological variation. This is important because patients are likely to have had previous medical intervention, were currently on a pharmaceutical programme, or being given treatment for a comorbidity or parallel indication, and therefore the effect of these states were to be investigated because they were the likely state of health of the recipients of stem cell therapy.

Diet, included tea intake, alcohol and tobacco intake (Figure 3) were included for similar reasons as the previous metrics, because they contained chemical agents that affect body chemistry, metabolism and health, and may have had a bearing on the quality of stem cells. They were included here to determine whether they had an adverse effect on biological variation.

BMI and weight are fundamental physical measurements. Physical measurements such as body fat and impedance were included as further, more detailed physical measurements. Pulse rate and blood pressure were included, alongside overall health rating, as direct physical measurements of a person's relative health, alongside health metrics such as fitness test duration and ECG (Figure 3). All available blood measurands were requested including cell counts, cell percentages and cell volumes amongst others, providing the potential for a direct comparison of how an 'unwell' population corresponds to the 'healthy' population, in respect to data spread and variation.

2.2 Statistical Analysis

The comma-separated variation output was processed in IBM's SPSS 22.0 statistics package. Nominal data fields, such as gender, were originally coded as numerical identities (1 = male, 2 = female). Unless stated otherwise, White Blood Cell count (WBC) was measured as cells / litre, noting that absolute cell numbers and cells per kilogram were unavailable in the UK Biobank database. Graphical representations including biological measurements such as WBCs were plotted on logarithmic scales to allow for visual inspection of spread and distribution that would otherwise be difficult on linear scales.

Distributions of cell metrics from the UK Biobank were analysed for normality, including graphical representation as a histogram. The minimum, maximum and range data were defined. Clearly, this is the most basic representation of the variation encountered, because these values represent the smallest and largest cell populations within a given time frame. They demonstrate the extremes of biological variation, and the potential extremes that a corresponding manufacturing process would need to be capable of addressing. Variance has been included because it is not based upon the mean (which may be skewed in non-parametric distributed data) but based on the distances from a given data point to another, so may be a more representative measure of spread in non-parametric distributed data.

These biological measurements were found to be non-parametric and consequently non-parametric statistical tools were used to compare median and distribution of the data. The Pearson product-moment correlation co-efficient was used to measure linear correlation. Independent samples median tests, independent Mann-Whitney tests (for two independent groups, such as gender) [16,17] and independent sample Kruskal-Wallis (for multiple independent groups) [18] tests were used to measure whether the median or the distribution of population differs significantly between groups. The Kruskal-Wallis (K-W) test is the non-parametric equivalent of the one-way ANOVA, that determines whether there are statistically significant differences

321 between two or more groups of an independent variable on a continuous or ordinal
322 dependent variable.
323

3.0 Results

The analysis of the UK Biobank data has concentrated on a range of metrics. Those defined for discussion (age, gender, geography, health state, weight) have been observed to be influential on the final variation statistics of the data, along with the general overview of the data set. Additional metrics such as smoking, alcohol, sleep, ethnicity, and exercise were analysed but were found to not have a significant influence on variation, and hence have not been reported here. Other potentially influential metrics such as disease state (beyond a statement for cancer), medication and diet have not been examined, but may be the subject of further analysis.

3.1 Overview of Variation

UK Biobank's haematology data includes a complete blood count although without CD marker quantification [15]. WBC count was used as an analogue in this work to examine biological variation in the relatively healthy population of the UK Biobank. WBCs were reported by UK Biobank as " $\times 10^9$ cells / litre" and as such, due to a lack of corresponding fields, absolute cell count, or cell count per kilogram of patient bodyweight (such as those reported in medical literature) was unavailable.

24,372 entries from UK Biobank contained missing data – no WBC data for example. At least 1,583 of those entries could not be measured for various reasons (e.g. the participants declined, equipment failure). Two entries were removed because the WBC count equalled zero, and would substantially skew the range of any reported variation. This brought the overall total to 478,290 records instead of the 502, 664 shown in Table 1.

Figure 5 is a scatter plot of WBC count per litre per individual measured, plotted against unique participant ID to illustrate the spread and distribution of this biological metric. The x-axis of the plot has been left blank because it is simply a linear list of participant identifiers. There are clear groupings between the regular expected spread, and outliers outside of this regular spread. This is suggestive of Parnaby's

manufacturing process descriptions of ‘regular runners’, ‘repeaters and ‘strangers’ [19]. The former two refer to product types a manufacturing facility is regularly exposed to and prepared for, whilst the latter represents dramatic changes in process. Within Figure 5 Runners and Repeaters is represented by the central band, where the majority of the donor characteristics lie – the band of variation that a theoretical processing machine or facility, with this starting material as its input (using WBC count as a surrogate measure), would be required to regularly handle. The samples outside of this band represent starting material the process may not be able to accommodate, and would potentially require adaptation or even a different process.

If this graph represents the biological variation of a potential starting material, then the ‘runners and repeaters’ marked on Figure 5 also represents the variation that elements of a particular process or automated system would have to manage or control – between approximately 2.4×10^9 cells / litre and 1.6×10^{10} cells / litre (Table 2), and between approximately 4.00×10^7 cells / litre and 3.90×10^{11} cells / litre irregularly (as indicated by ‘strangers and aliens’ variation). Hence, a variation of four orders of magnitude around the median (6.65×10^9 cells / litre). It is noted here that there may be only a few, or many points in a particular biomanufacturing process that could be influenced by such variation, but any potential examples would be cell therapy specific.

		White Blood Cell Count (WBC)						
	Participants (n)	Mean	Median	Minimum	Maximum	Range	Standard Deviation	Variance
Dataset	478,290	6.89×10^9	6.65×10^9	4.00×10^7	3.90×10^{11}	3.90×10^{11}	2.12×10^9	4.51×10^{18}
Male	219,317	6.92×10^9	6.69×10^9	8.00×10^7	1.90×10^{11}	1.89×10^{11}	2.19×10^9	4.81×10^{18}
Female	258,958	6.86×10^9	6.62×10^9	4.00×10^7	3.90×10^{11}	3.90×10^{11}	2.06×10^9	4.25×10^{18}
Has not Had Cancer	439,015	6.88×10^9	6.66×10^9	4.00×10^7	1.04×10^{11}	1.04×10^{11}	1.82×10^9	3.30×10^{18}
Has / Has Had Cancer	39,275	6.99×10^9	6.58×10^9	1.00×10^9	3.90×10^{11}	3.90×10^{11}	4.24×10^9	1.80×10^{19}

Table 2: WBC variation statistics defined within the population (cells / litre)

378

379 **3.2 Variation and outliers as a function of age**

380 The population of UK Biobank is weighted towards a more mature population (40 –
381 69) than the previously published Meta-analysis dataset, which included paediatric
382 patients and donors [8, 9]. Due to the increase in life expectancy and the subsequent
383 growth of the ageing population, this demographic is important, particular as age
384 related diseases and dysfunctions may be a key target for regenerative medicine,
385 and better mirrors a recently completed single clinical centre analysis of starting
386 material variation [11].

387 WBC count / litre relative to age in years is shown in Table 3 and Figure 6. Age
388 appears to initially have an insignificant effect on the variation of WBCs, but there is

a marked increase in the number of extreme values and hence range from 60 years of age and onwards highlighted in the dashed line box in Figure 6. This particular spike in 'outliers' and the subsequent increase in the spread of cell counts may be a result of the diverse ways in which people age relative to one another. The mean and core values (represented by the one standard deviation) remain relatively constant across age groups. The significant increase in outliers as the patient's age increases may be due to the increased likelihood of the effects of diagnosed or undiagnosed age-related diseases within the participants. It is noted that current literature identifies that typically WBC count declines with age, but increases with disease indicators [20]. There are similarities here to the single clinical centre analysis [11].

Age (Years)	Participants (n)	White Blood Cell Count (WBC)						
		Mean	Median	Minimum	Maximum	Range	Standard Deviation	Variance
<45	2	7.44x10 ⁹	7.44x10 ⁹	6.63x10 ⁹	8.24x10 ⁹	1.61x10 ⁹	1.14x10 ⁹	1.30x10 ¹⁸
46 – 50	26,942	6.89x10 ⁹	6.67x10 ⁹	9.00x10 ⁷	1.93x10 ¹⁰	1.92x10 ¹⁰	1.83x10 ⁹	3.35x10 ¹⁸
51 – 55	59,631	6.88x10 ⁹	6.65x10 ⁹	1.90x10 ⁸	8.87x10 ¹⁰	8.85x10 ¹⁰	1.88x10 ⁹	3.53x10 ¹⁸
56 – 60	68,698	6.83x10 ⁹	6.60x10 ⁹	1.00x10 ⁹	3.90x10 ¹¹	3.89x10 ¹⁰	2.48x10 ⁹	6.13x10 ¹⁸
61 – 65	80,089	6.78x10 ⁹	6.55x10 ⁹	8.00x10 ⁷	1.28x10 ¹¹	1.27x10 ¹¹	1.98x10 ⁹	3.91x10 ¹⁸
66 – 70	111,857	6.85x10 ⁹	6.60x10 ⁹	1.30x10 ⁸	1.13x10 ¹¹	1.13x10 ¹¹	2.02x10 ⁹	4.09x10 ¹⁸
71 – 75	98,530	7.00x10 ⁹	6.77x10 ⁹	4.00x10 ⁷	1.90x10 ¹¹	1.89x10 ¹¹	2.30x10 ⁹	5.29x10 ¹⁸
76 – 80	32,525	7.04x10 ⁹	6.80x10 ⁹	9.80x10 ⁸	1.04x10 ¹¹	1.02x10 ¹¹	2.06x10 ⁹	4.24x10 ¹⁸
81+	1	8.00x10 ⁹	8.00x10 ⁹	8.00x10 ⁹	8.00x10 ⁹	0	-	-

Table 3: WBC count for participants stratified into age groups of 5 years (cells / litre)

3.3 Variation and outliers as a function of health state

The participants of the UK Biobank survey are notionally healthy, a baseline of the UK adult population – however there are participants who have had or have serious illness such as cancer, hence it was important to stratify these participants and compare the relatively healthy population with that of the sick population. This has particular potential when considering investigating correlations with single site clinical centres which would be almost exclusively sick populations.

In this example, 'healthy' participants were identified as those who had declared being cancer-free, and 'un-healthy' participants were identified as those who had declared that they had had, or were currently suffering from cancer (Table 2). The aim was to identify whether there was a significant difference between a healthy and an un-healthy population, noting that it is well understood that WBC will change as a function of a variety of cancers. However, it should be identified that within the UK Biobank participant questionnaire there was very limited questioning about potential

previous or existing cancer tests (bowel, prostate, breast), and a general question about GP informing the presence of cancer. Consequently, there was not enough resolution of information within the UK Biobank dataset to extend the investigation to consider type/stage of cancer, or treatments being used.

Figure 7 shows comparative plots of healthy (439,015) and un-healthy (39,275) participants with respect to cancer. The x-axis of both elements of the plot has been left blank because it is simply a linear list of participant identifiers. Variation as a function of range between individuals is higher for those individuals who have had cancer compared to those who have not. In addition, for those who have had cancer, the predominant variation is significant number of data points showing an increase of WBC beyond the mean value, whilst for those who have not had cancer, there is minimal variation (in terms of absolute number of data points) around the mean, albeit with a larger range. This is probably due to the wide variety of results covered by 'has or had cancer' in terms of type, severity, longevity, treatment type and whether the patient is in remission or relapse, compared to a comparatively stable, ambulatory population.

However, this does not take into account the differences between cancer patients – when did the illness occur, severity, treatment, remission – this example has assumed that exposure at some level to a malignancy will have a noticeable effect on biological variation as a function of WBCs. An extension of this investigation would be the variation as a function of cell quality and function with respect to cancer stage and type – something that is not possible with this dataset, but may be a more appropriate measure of variation. Interestingly, the mean and median values are very similar between the two groups (approximately 6.90×10^9 cells / litre and 6.60×10^9 cells / litre respectively).

3.4 Variation and outliers as a function of gender

UK Biobank originally contained 229,182 male participants and 273,467 female participants. Due to missing data as defined in Section 3.1, WBC count was only recorded for 219,317 male and 258,958 female donors. WBC count / litre relative to gender has been plotted in Figure 8. The range of WBC counts for male participants was between 8.00×10^7 cells / litre to 1.90×10^{11} cells / litre, and for female participants was between 4.00×10^7 cells / litre and 3.90×10^{11} cells / litre. The median cells / litre is higher for male participants than female patients (Table 2).

There is an increased range for female participants over male participants (3.90×10^{11} cells / litre versus 1.89×10^{11} cells / litre), and therefore overall variation from a biomanufacturing perspective as a starting material. However, through visual inspection of the box and whisker plot (Figure 8) this may be due to a small number of extreme cases. Both the median and the distributions are significantly different between genders ($p < 0.05$) indicating that genders may be distinct sub-populations with respect to WBC, and that variation may be different between the two. For the whole population of UK Biobank, the variance between participants is higher for female participants – meaning the product to product variation in terms of cell number will be higher – but the range of male participant's WBC counts are higher. Consequently, this analysis suggests that gender of the donor may have a significant effect on biomanufacturing process design.

--

475

476

477 3.5 Variation and outliers as a function of weight

478 Weight is a key physical metric for determining biological measurements such as
479 blood. As a participant's physical mass increases, typically so will the volume of
480 blood. Weight is a commonly used metric in medicine, used to determine
481 pharmaceutical dose for example. The weight of a given participant in the UK
482 Biobank was 78.1 kg (mean) and 76.4 kg (median) ($n = 499,874$). Table 4 shows the
483 WBC count for binned weight categories whilst Figure 9 is the box and whisker plot
484 of WBC count against corresponding weight, where weight has been categorised
485 into 10 kg bins. In both the overall population and a representative, randomly chosen
486 sample of the population (20%) these groups are significantly different in terms of
487 both median WBC count and distribution of cell counts ($p < 0.05$ for both). This is
488 unsurprising because the median cell count increases with participant weight.

489

490

491

492

493

494

495

496

497

498

499

500

Weight (kg)	Participants (n)	White Blood Cell Count (WBC)						
		Mean	Median	Minimum	Maximum	Range	Standard Deviation	Variance
<30	1	7.03x10 ⁹	7.03x10 ⁹	7.03x10 ⁹	7.03x10 ⁹	0	-	-
30.1 – 40.0	111	6.50x10 ⁹	6.28x10 ⁹	2.86x10 ⁹	1.23x10 ¹⁰	9.46x10 ⁹	1.90x10 ⁹	3.62x10 ¹⁸
40.1 – 50.0	5,461	6.71x10 ⁹	6.42x10 ⁹	1.59x10 ⁹	2.52x10 ¹⁰	2.36x10 ¹⁰	1.94x10 ⁹	3.77x10 ¹⁸
50.1 – 60.0	48,773	6.64x10 ⁹	6.40x10 ⁹	8.00x10 ⁸	1.25x10 ¹¹	1.14x10 ¹¹	2.05x10 ⁹	4.21x10 ¹⁸
60.1 – 70.0	106,876	6.72x10 ⁹	6.49x10 ⁹	9.00x10 ⁷	3.90x10 ¹¹	3.90x10 ¹¹	2.42x10 ⁹	5.84x10 ¹⁸
70.1 – 80.0	119,627	6.82x10 ⁹	6.60x10 ⁹	4.00x10 ⁷	1.07x10 ¹¹	1.07x10 ¹¹	1.98x10 ⁹	3.93x10 ¹⁸
80.1 – 90.0	97,825	6.94x10 ⁹	6.70x10 ⁹	6.40x10 ⁸	1.81x10 ¹¹	1.81x10 ¹¹	2.04x10 ⁹	4.15x10 ¹⁸
90.1 – 100.0	55,698	7.08x10 ⁹	6.87x10 ⁹	7.00x10 ⁷	1.46x10 ¹¹	1.46x10 ¹¹	2.06x10 ⁹	4.23x10 ¹⁸
100.1 – 110.0	25,480	7.27x10 ⁹	7.06x10 ⁹	1.90x10 ⁸	8.08x10 ¹⁰	8.06x10 ¹⁰	1.97x10 ⁹	3.89x10 ¹⁸
110.1 – 120.0	10,166	7.50x10 ⁹	7.30x10 ⁹	9.80x10 ⁸	4.06x10 ¹⁰	3.97x10 ¹⁰	1.96x10 ⁹	3.82x10 ¹⁸
120.1 – 130.0	4,018	7.69x10 ⁹	7.50x10 ⁹	2.20x10 ⁹	2.25x10 ¹⁰	2.03x10 ¹⁰	1.91x10 ⁹	3.63x10 ¹⁸
130.1 – 140.0	1,538	7.93x10 ⁹	7.71x10 ⁹	3.10x10 ⁹	1.97x10 ¹⁰	1.66x10 ¹⁰	1.97x10 ⁹	3.89x10 ¹⁸
140.1 – 150.0	658	7.98x10 ⁹	7.90x10 ⁹	2.80x10 ⁹	1.67x10 ¹⁰	1.39x10 ¹⁰	1.82x10 ⁹	3.32x10 ¹⁸
150.1 – 160.0	253	8.23x10 ⁹	8.01x10 ⁹	2.30x10 ⁹	1.43x10 ¹⁰	1.20x10 ¹⁰	1.89x10 ⁹	3.56x10 ¹⁸
160.1 – 170.0	113	7.91x10 ⁹	7.61x10 ⁹	4.28x10 ⁹	2.07x10 ¹⁰	1.64x10 ¹⁰	2.47x10 ⁹	6.12x10 ¹⁸
170.1 – 180.0	58	8.31x10 ⁹	8.33x10 ⁹	4.20x10 ⁹	1.85x10 ¹⁰	1.43x10 ¹⁰	2.29x10 ⁹	5.23x10 ¹⁸
180.1 – 190.0	13	8.34x10 ⁹	8.51x10 ⁹	6.40x10 ⁹	1.19x10 ¹⁰	5.50x10 ⁹	1.66x10 ⁹	2.76x10 ¹⁸
190.1+	7	9.03x10 ⁹	8.10x10 ⁹	5.20x10 ⁹	1.26x10 ¹⁰	7.30x10 ⁹	2.86x10 ⁹	8.16x10 ¹⁸

Table 4: WBC count for participants stratified into binned weight groups of 10 kg (cells / litre)

The spread of the data does not appear to be correlated to weight. Variation and spread in the data tends to tail off towards the heavier weights; this is likely because of the variety of body types that can be encapsulated within lighter weights (shorter and heavier versus taller and thinner for example) compared to the heavier weights which are likely to be more limiting in body shape (and therefore variation in the amount of blood required to supply). This effect may well be dependent on the significant clustering around the median value, where most of the data occurs. Whilst the metrics available in the UK Biobank dataset at the time of analysis did not allow for better determination of cause and effect with respect to weight, this issue is important because cell therapy dose is currently often defined as a function of patient

weight. In addition, significant work is already available that identifies links between WBC count and obesity [21].

Figure 9 is particularly interesting from a variation and outliers perspective, because the median WBC count appears to increase with patient size, and the interquartile ranges appear to be comparable between weight categories, but the outliers and the far outliers (circles and crosses) have considerable impact on the expected spread of WBC counts for a given starting material for each category.

For a given weight category (such as the modal category 70.1 kg to 80.0 kg), the number of 'outliers' can be calculated. For data points to be considered outliers in the box and whisker plot, they need to be outside of 1.5 times the interquartile range (the width of the 'box') above the 25th quartile or the 75th quartile (the brackets of the box and whisker plot). As a result, for the 70.1 kg to 80.0 kg category there are 1,257 outliers, which account for 1.05% of the total category. Without these outliers the range of WBC count within this category would be between 2.29×10^9 cells / litre and 1.12×10^{10} cells / litre, compared to 4.00×10^7 cells / litre to 1.07×10^{11} cells / litre when including these outliers.

If it could be identified whether there are factors that contribute to these 'outliers' or whether they are part of the naturally occurring biological variation is an important issue because it would determine whether a manufacturing process could work within the boundaries of a box and whisker plot, or whether these outliers must be considered. These outliers are a small proportion of the overall population but cannot be easily discounted (i.e. 1,257 people) but they have a strong influence on the potential integrity of a manufacturing process.

3.6 Variation and outliers as a function of location of Assessment Centre

Assessment Centre (across the UK) and attending date were included in the research definition, to examine whether biological variation was affected by the geographical location of different sub-populations. These different sub-populations were likely to contain localised ratios of particular social class or ethnic groups – which is one of the reasons why blood banks are typically set up in particular areas of the UK.

Figure 10 is a box and whisker plot illustrating the spread of WBC count / litre between separate locations categorised by assessment centre. According to

independent samples median test ($H(20) = 1,893.425$, $p < 0.05$) and independent samples Kruskal-Wallis test ($H(20) = 2,702.357$, $p < 0.05$), the median cell count and distributions of cell counts were significantly different between centres. Potentially, the cause for these differences in median or spread of cell counts is a function of geographical location. This could be local environmental conditions such as quality of life, air or water quality, cultural influences, or accessibility to health amenities, although a more extended analysis specifically looking at geographical influences would need to be completed to give better determination. An alternative reason could be procedural differences at the collection centres, or the transit time and methodology – although UK Biobank has put considerable effort into ensuring the equivalency of process and protocols between centres, and analysis of all samples in a centralised location.

Comparing the variation between geographical centres (Figure 10) allows for further illustration of the compounding effect of patients with ‘extreme’ cell counts. Taking the most populated category (Leeds, $n = 44,215$) as an example, and by counting those values that exceed 1.5 times the interquartile range outside of the 25th and 75th quartile, 5.5% (2,330) of this sub-population could be identified as outliers. Without these outliers the range of cell counts within the Leeds catchment area would be between 3.45×10^9 cells / litre and 9.90×10^9 cells / litre. Including these outliers (but excluding the one zero value) the range of cell counts is between 8.00×10^8 cells / litre and 8.68×10^{10} cells / litre (i.e. 2 orders of magnitude). The question remains whether these ‘outliers’ are part of the natural background variation (common cause variation), or are as a result of a particular influence or factor (special cause variation) and in which case, this factor or factors will need to be identified, this being critical for understanding the behaviour of the starting material and the process.

4.0 Conclusions

The scale up of autologous and allogeneic regenerative medicine products from clinically-based to centralised or decentralised manufacturing facilities is the subject of much current discussion and debate. The viability of any scaled-up therapeutic process, or indeed any viable manufacturing process, will be dependent on an understanding of the potential variation of the starting materials and whether the manufacturing process can accommodate or control such variation.

The aim of this research has been to quantify the White Blood Cell (WBC) count variation encountered in a large-scale nominally healthy population, represented by the UK Biobank, and driven by previous work that has examined the variation of CD34+ starting materials in HSCT. This new analysis has been an opportunity to examine and benchmark the baseline biological variation (as expressed by WBC count) in the general population and whether there are any contributing or confounding factors. This information is needed to inform future sources of (specifically) allogeneic starting materials, and to help inform the design of cell therapy processing, manufacturing and process control.

It has been found that the variation in WBC count can be between 7.00×10^7 cells/litre and 3.90×10^{11} cells/litre – a variation of up to four orders of magnitude. Both the Total Nucleated Cell (TNC) count as defined by previous sick population analysis [8,9] and these current WBC counts within the UK Biobank have similar ranges (range; 3.90×10^{11} and 3.86×10^{11} cells / litre), similar distributions around the mean (standard deviation; 2.12×10^9 cells / litre and 2.10×10^9 cells / litre), and similar data spread (variance; 4.51×10^{18} cells / litre and 4.42×10^{18} cells / litre), respectively.

An important assumption that has been made here is that the UK Biobank is a representative sample of the UK general population. The original data gathering studies that form the UK Biobank were carried out on a volunteer-basis that may

611 have influenced the population, for instance people of a mindset that were more
612 inclined to volunteer, and may reflect a particularly health state, lifestyle or character.

613

614 At the time of analysis, the UK Biobank had no direct measure of stem cells identity
615 or stem cell quality, such as the CD34+ marker. It has been proposed here that WBC
616 count is an appropriate illustrative analogue for the potential variation in starting
617 materials for cell therapies, with WBC count varying as a function of key biophysical
618 metrics. Consequently, there has been no claim or attempt to establish or form any
619 direct correlation with other previous studies that have concentrated on the CD34+
620 marker, although we have chosen to draw parallels between the two studies for
621 illustrative purposes.

622

623 Furthermore, it is not possible to make specific inference or correlation to the quality
624 or variation of stem cells derived from peripheral blood or bone marrow aspiration,
625 given that all UK Biobank participants donated peripheral whole blood samples.
626 Likewise, this work has been driven on the basis of expertise in biomanufacturing
627 and biometrology and the current remit does not include any correlations with
628 dosage, clinical efficiency or clinical justification for WBC count variation. This would
629 be the subject of significant future work based on an exemplar cell therapy process.

630

631 Whilst our analysis of WBC count variation could be perceived to be not of initial
632 direct relevance to autologous therapies (whereby the patient is the source of their
633 own stem cells), in some cases a patient may need to transition from autologous to
634 allogeneic therapy starting material as a function of being too ill to provide sufficient
635 quantities of their own cells. Likewise, allogeneic therapies can potentially take
636 opportunity to pick from a donor that would perform best in a biomanufacturing
637 process. However, if tissue typing, rejection or GVHD is the initial concern, then
638 optimisation of other donor metrics may end up being subsidiary decision points, and
639 thus our analysis again provides a valuable viewpoint.

640

641 The population characteristics examined in this research (age, disease state,
642 gender, geographical location and weight) all appear to influence the biological
643 variation encountered but there is inherently an implication this variation relates to a
644 multiplicative effect based on multiple, interacting variables, rather than necessarily
645 the result of individual or a handful of key metrics. There are several potentially
646 influential metrics that have not been specifically discussed in this work which may
647 play a role e.g. smoking, alcohol, sleep, ethnicity, and exercise. These metrics were
648 analysed during the research but were found to not have a significant influence on
649 variation compared to the reported metrics, and hence have not been detailed here

650

651 To summarise the observed influences of these characteristics:

- 652 • Gender appears to impact the amount of variation in WBC and TNC counts
- 653 • Increasing age shows a trend towards increasing variation in WBC and TNC
654 counts
- 655 • Increasing weight shows a trend towards decreasing variation in WBC and
656 TNC counts
- 657 • UK Assessment Centre, analogous of geographical area, appears to have a
658 statistically significant effect on WBC counts

659

660 A notable observation from this analysis is the considerable number of data 'outliers'
661 and the non-parametric distributions illustrated by the data analysis. Outliers in
662 traditional manufacturing would represent material or product to be discarded as out-
663 of-specification; however, in cell therapy they are too valuable and too unique to be
664 discarded. These outliers represent a state of health, keyed uniquely to be the
665 individual, and cannot be removed to smoothen a statistical distribution.

666 Understanding these outliers, and what they represent, is critical to understanding
667 the behaviour of the starting material and the process; it will define the way in which
668 cellular therapies are manufactured and monitored, and act as a guide as to whether
669 biological variation should be accommodated or could be controlled.

Summary Points

- Cell therapies are increasingly viable for upscale to centralised or distributed manufacture becoming possible. Significant variation of starting materials in HSCT has been previously identified, and there is a need to examine a healthy population for variation in blood metrics to inform potential manufacturing process design and control paradigms.
- UK Biobank is a unique database of over 500,000 volunteer participants available for medical exploration, that has in this research been made available for variation analysis of blood metrics.
- White Blood Cell count has been used as an analogue of variation for cell therapy input/starting material, in comparison to previous community wide and single site clinical center studies examining variation of starting material in HSCT.
- Overall, White Blood Cell variation of four orders of magnitude around the mean (6.65×10^9 cells / litre) was found.
- Distributions were predominantly found to be asymmetric and non-parametric with distribution characteristics being statistically robust because of the sample size.
- Key correlates were investigated. Variation increased with age, and increased with weight up to 80 kg then decreased. Health state showed a greater absolute number of participants with elevated count. Female range was found to be greater than male. Cell count / distributions were found to be different between UK Biobank collection centres.

698

699

- This research identifies that variation will not be uniquely related to one cause, but will be a function of multiple factors. This magnitude of variation signals a requirement for new strategies for manufacturing process design and control that are either; able to manage such ranges of process input, or the variation found in starting materials should be better controlled.

700

701

702

703

704

705

706 **Translational Perspective**

707

708 The work reported in this research has determined the potential for variation in blood
709 metrics within the UK Biobank database of 500,000+ nominally healthy participants.
710 Variation as a function of age, gender, weight, health state, and/or collection centre
711 is seen to be up to 4 orders of magnitude as a function of White Blood Cell (WBC)
712 count measured in cells / litre. It is shown that this variation has similarities to
713 variation previously seen in starting materials for Haematopoietic Stem Cells
714 Therapy (HSCT) exemplars, both on a global population basis and at a specialist
715 single clinical centre.

716

717 If cell therapies are to be continually and successfully translated out of clinical single
718 dose production to larger scale (centralised or distributed) biomanufacture, then one
719 important aspect of process design (amongst many) is to be able to accommodate or
720 control any variation of the starting materials. But to do this requires an informed
721 understanding of the potential of variation that may exist within typical starting
722 materials. This analysis of UK Biobank has achieved this for a statistically significant
723 nominally healthy population in terms of WBC count, given that at the time of the
724 analysis the database did not contain any further cell identity marker data (e.g
725 CD34+).

726

727 The WBC count variation shown in this analysis illustrates the potential for blood
728 metric variation within a potential donor pool, variation which most existing
729 manufacturing process control systems are not designed to accommodate and
730 likewise would not be able to control. Hence this work provides opportunities to
731 further consider the design of such biomanufacturing process control systems. In
732 addition, it informs the process of determining if the initial variation of cell therapy
733 starting materials requires better management and control prior to the
734 biomanufacturing process start, or, whether the biomanufacturing process controls
735 need to be redesigned to accommodate any potential or existing variation.

736 **Ethical Conduct of the Research**

737 Blood metric data was supplied by UK Biobank. All ethical requirements for use of
738 such material had been completed by UK Biobank during the collection stage of the
739 UK Biobank project.

740

741 **Acknowledgements**

742 This research is kindly supported by the UK EPSRC/MRC Doctoral Training Centre
743 for Regenerative Medicine at Loughborough University (EP/L105072/1), and UK
744 Biobank for access to UK biometric datasets (Application No. 4047).

745

References

- 1 Hourd P, Chandra A, Medcalf N, Williams DJ. Regulatory challenges for the manufacture and scale-out of autologous cell therapies, StemBook, ed. The Stem Cell Research Community, StemBook, 2014, doi/10.3824/stembook.1.96.1.
- 2 Hourd P, Williams DJ. Scanning the horizon for value-add manufacturing science: Accelerating manufacturing readiness for the next generation of disruptive, high-value curative cell therapeutics. *Cytotherapy*. 20, 759-767. (2018).
- 3 EudraLex, The Rules Governing Medicinal Products in the European Union, Volume 4, Good Manufacturing Practice, Guidelines of 22.11.2017: Good Manufacturing Practice for Advanced Therapy Medicinal Products, European Commission, Brussels, 22.11.2017 C(2017) 7694 Final, 2017.
- 4 Marks P, Gottlieb S. Balancing safety and innovation for cell-based regenerative medicine, *N Engl J Med*. 378(10), 954-959 (2018).
- 5 Harrison RP, Rafiq QA, Medcalf N. Automating decentralized manufacturing of cell and gene therapy products. *Cell Gene Therapy Insights*. 2(4), 489-497 (2016).

- 771 6 Harrison RP, Ruck S, Medcalf N, Rafiq QA. Decentralized manufacturing of
772 cell and gene therapies: Overcoming challenges and identifying opportunities,
773 Cytotherapy. 19,1140-1151 (2017).
774
- 775 7 Harrison RP, Rafiq QA, Medcalf N. Centralised versus decentralized
776 manufacturing and the delivery of healthcare products: A United Kingdom
777 exemplar. Cytotherapy. 20, 873-890 (2018).
778
- 779 8 Thurman-Newell JA, Petzing J, Williams D. Quantification of biological
780 variation in blood-based therapies: A summary of a meta-analysis to inform
781 manufacturing in the clinic, Vox Sanguinis. 109(4), 394-402 (2015).
782
- 783 9 Thurman-Newell JA, Petzing J, Williams D. A meta-analysis of biological
784 variation in blood-based therapy as a precursor to biomanufacturing.
785 Cytotherapy. 18, 686-694 (2016).
786
- 787 10 Thurman-Newell JA. Quantification of variation in biological input materials
788 and its effect on manufacture and process control. PhD Thesis,
789 Loughborough University, UK, 2017, [https://dspace.lboro.ac.uk/dspace-](https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/25346/3/Thesis-2017-ThurmanNewell.pdf)
790 [jspui/bitstream/2134/25346/3/Thesis-2017-ThurmanNewell.pdf](https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/25346/3/Thesis-2017-ThurmanNewell.pdf)
791
- 792 11 Thurman-Newell JA, Petzing JN, Williams DJ. Analysis of biological variation
793 in blood-based therapy as a precursor to biomanufacturing, MIT Center for
794 Biomedical Innovation – 2015 BioMAN Summit, Boston, USA, 18th – 19th
795 (2015).
796
- 797 12 UK BioBank, <http://www.ukbiobank.ac.uk>, 2018 (accessed 02 October 2018).

798

799 13 Lee CD, Folsom AR, Nieto FJ, Chambless LE, Shahar E, Wolfe DA. White
800 blood cell count and incidence of coronary heart disease and ischemic stroke
801 and mortality from cardiovascular disease in African-American and white men
802 and women: Atherosclerosis risk in communities study, Am. J. Epidemiol,
803 154(8), 758–764 (2001).

804

805 14 Allen N, Sudlow C, Downey P, Peakman T, Danesh J, Elliott P, et al. UK
806 Biobank: Current status and what it means for epidemiology, Health Policy &
807 Technology 1, 123–126 (2012).

808

809 15 UK Biobank: Protocol for a large-scale prospective epidemiological resource
810 Protocol No: UKBB-PROT-09-06 (Main Phase), 2007, (accessed 25 June
811 2019)

812

813 16 Mann HB, Whitney DR. On a test of whether one of two random variables is
814 stochastically larger than the other. Ann. Math. Stat. 18(1), 50-60 (1947).

815

816 17 Fay MP, Proschan MA. Wilcoxon–Mann–Whitney or t -test? On assumptions
817 for hypothesis tests and multiple interpretations of decision rules. Statistics
818 Surveys 4, 1–39 (2010).

819

820 18 Kruskal W, Wallis A. Use of ranks in one-criterion variance analysis. J. Am.
821 Stat. Assoc. 47(206), 583–621 (1952).

822

823

824

- 825 19 Parnaby J. A systems approach to the implementation of JIT methodologies in
826 Lucas Industries. *Int. J. Prod. Res.* 26(3), 483–492 (1988).
827
- 828 20 Zohreh A, Parsa E. Relationship between age and peripheral white blood cell
829 count in patients with sepsis. *Int. J. Prev. Med.* 2(4), 238-242 (2011).
830
- 831 21 Dixon JB, O' Brien PE. Obesity and the white blood cell count: Changes with
832 sustained weight loss, *Obesity Surgery*, 16(3), 251-257 (2006).
833

834 **List of Tables**

835 Table 1: Summary of participants within UK Biobank [12]

836

837 Table 2: WBC variation statistics defined within the population (cells / litre)

838

839 Table 3: WBC count for participants stratified into age groups of 5 years (cells / litre)

840

841 Table 4: WBC count for participants stratified into binned weight groups of 10 kg
842 (cells / litre)

843

844 **List of Figures**

845 Figure 1: Participant Metrics – General, Ethnicity and Medical Health

846

847 Figure 2: Participant Metrics – Physical Body Measurements

848

849 Figure 3: Participant Metrics – Diet and Exercise

850

851 Figure 4: Blood Composition Metrics

852

853 Figure 5: Variation of WBC measured per individual participant (cells / litre)

854

855 Figure 6: The spread of WBC count per individual participant as a function of age
856 (cells / litre)

857

858 Figure 7: WBC count per individual participant that have had or have cancer versus
859 those who have not (cells / litre)

860

861 Figure 8: WBC count against gender (cells / litre)

862

863 Figure 9: Calculated WBC count per individual as a function of weight in kilograms
864 (cells / litre)

865

866 Figure 10: WBC count per individual as a function of location of UK Assessment
867 Centre (cells / litre)