

Title: ‘Cervified’: a method for the morphometric identification of red deer (*Cervus elaphus*), fallow deer (*Dama dama*), and roe deer (*Capreolus capreolus*) archaeological bones.

Journal name: Journal of Archaeological Method and Theory

Authors: Veronica Aniceti^{1*}, Mauro Rizzetto^{2*}, Francesco Giacalone³

¹ Corresponding author. Consejo Superior de Investigaciones Científicas, Institució Milà i Fontanals (CSIC-IMF), Barcelona, Spain; veronica.aniceti@imf.csic.es

² Corresponding author. Vrije Universiteit Brussel, Interdisciplinary Historical Food Studies research group (VUB-FOST), Brussels, Belgium; mauro.rizzetto@vub.be

³ Ludwig.guru, Palermo, Italy

* These authors contributed equally to this work.

Online Resource 2

Inter-observer and intra-observer error assessments. Description of the method employed for the assessments of the inter- and intra-observer errors, description and discussion of results.

Measurement reliability was assessed through both inter-observer and intra-observer agreement analyses using the Intraclass Correlation Coefficient (ICC). All statistical analyses were performed using the R statistical software, with the 'irr' and 'psych' packages.

Inter-observer reliability was quantified using a two-way random-effects model (ICC type 2,1) with absolute agreement definition. This model was selected because multiple observers, considered representative of a larger population of potential observers, evaluated each specimen once [35].

Intra-observer reliability was assessed using a one-way random-effects model (ICC type 1,1). This model is appropriate when each specimen is measured multiple times (i.e., three times across as many days) by a single observer, who is considered a random representative of a larger population of potential observers [35]. In the specific case of the Cervified dataset, all measurements were consistently taken by the same two researchers (Veronica Aniceti and Mauro Rizzetto). Given the high level of standardisation in data recording, ensured by the measurement protocol and the close coordination between these researchers, they were considered as a single observer for analytical purposes.

ICC values were interpreted according to established guidelines in statistical literature: values below 0.50 indicate poor reliability, those between 0.50-0.75 moderate reliability, between 0.75-0.90 good reliability, and values above 0.90 indicate excellent reliability. For all analyses, 95% confidence intervals were calculated; statistical significance was established at $p < 0.05$, with p testing the null hypothesis that $ICC = 0$ [35].

In sum, an ICC value closer to 1 indicates higher agreement between measurements, showing that different observers (inter-observer) or the same observer at different times (intra-observer) are obtaining very similar measurements. An ICC value closer to 0, on the other hand, suggests poor agreement and high variability in the measurements. The confidence intervals (Lower CI and Upper CI) show the range within which the true ICC value is likely to fall with 95% confidence. Narrow intervals suggest more precise estimates.

The analysis of inter-observer error demonstrates remarkably strong consistency across most anatomical elements and measurements (S2 Table 1). With 77.8% of all measurement tests showing statistical significance ($p < 0.05$) and 41.3% achieving high significance ($p < 0.001$), these results validate the robustness of the recording protocol employed. This is a consequence of the high ICC

values (predominantly > 0.90 , indicating excellent reliability) across virtually all measurements, which reinforce the reliability of the methodological framework of this study.

An intriguing pattern emerges from the analysis of p values across different measurement types: approximately 13 'traditional' measurements show p values ≥ 0.05 (the GLs of metacarpus and metatarsus, Bd and BFd on various elements, etc.). The reason behind the low performance of these routinely taken measurement is rather unusual. This result is probably due to one or more errors by the observers that is probably not significant for our understanding of the replicability of that measurement. Despite the relatively higher p values, however, these measurements still produce remarkably high ICC values (most of them > 0.90 , i.e. excellent reliability). The ICC values for the intra-observer error across all measurements are remarkably high, ranging from 0.994 to 0.999, with narrow confidence intervals (S2 Table 2). These exceptionally high values indicate an excellent intra-observer reliability for all measurements. All tests show highly significant results, with $p < 0.001$. Such high intra-observer reliability suggests that measurement error is negligible. This strengthens the foundation for any interpretations derived from the use of such measurements, as the differences observed between specimens would predominantly reflect genuine size and morphological variations rather than measurement inconsistency.

Element	Measurement	ICC	Lower CI	Upper CI	<i>p</i> value
Scapula	SLC	0.989783705046606	0.954549351360824	0.999738784966897	< 0.001
	GLP	0.99866748114334	0.993942160598559	0.99996617783169	< 0.001
	LG	0.954325282317673	0.810084476889104	0.998792788826213	< 0.001
	BG	0.99932332255123	0.996913790208992	0.999982833809318	< 0.001
	ASG	0.96995417382165	0.874426231006572	0.999218394043706	0.0055
Humerus	GLC	0.981989594913997	0.890342564847681	0.999981259004102	≥ 0.05 (0.1419)
	SD	0.999268663263041	0.996643647668369	0.999981445984981	0.0122
	SD1	0.984992487265456	0.934560298039838	0.999614652814361	≥ 0.05 (0.0804)
	BT	0.982241763385676	0.888635208742605	0.999981390065965	≥ 0.05 (0.1440)
	Bd	0.992970326614785	0.946219633812944	0.999992542074256	0.0046
	HTC	0.88998884970589	0.517704132756425	0.999875839834684	0.0130
	HX	0.985932317952923	0.937358249912422	0.999638888376615	0.0347
	HTL	0.996664385673925	0.977996258299494	0.999996562719093	0.0348
	GL	0.879319482258689	0.498941947547669	0.999862084264253	≥ 0.05 (0.1918)
	SD	0.977535165606657	0.901669591291933	0.999419890912198	0.0043
Radius	SD1	0.850786205504714	0.428563506431617	0.999824809107905	< 0.001
	Bp	0.995332546967468	0.969981588314217	0.999995199111786	< 0.001
	BFp	0.997100178507056	0.986794701596034	0.999926293288842	< 0.001
	Dp	0.980726927095675	0.915716451254175	0.99950302364699	0.0455
	V	0.937664510386423	0.670406720779153	0.999932030470427	0.0027
	GL	0.967858653781925	0.807758131137978	0.999965898694036	0.0037
	SD	0.992459719354957	0.953666417437047	0.999992260036549	≥ 0.05 (0.1255)
Metacarpus	SD1	0.965719773204752	0.813169308995413	0.999964036070827	0.0017
	Bp	0.979155875584947	0.875889632905716	0.999978300595722	0.0375
	Dp	0.739562708952088	0.243366312583592	0.999658776731811	0.0117
	BFd	0.936201776527609	0.67757847044806	0.999931083498054	0.0268
	a	0.992719233065239	0.954996499707036	0.999992528403814	≥ 0.05 (0.1209)
	b	0.985409480222964	0.913690367020315	0.999984944459396	0.0406
	l	0.976462921122177	0.865023237777938	0.999975524371048	≥ 0.05 (0.3628)

	2	0.997578828693928	0.984252987446517	0.999997514905291	≥ 0.05 (0.0649)
	3	0.996873324172644	0.980471265718508	0.999996806385777	0.0122
	4	0.932144801558352	0.675550671775069	0.999926461240033	≥ 0.05 (0.0666)
	5	0.998369594008432	0.987624036416221	0.999998291013257	< 0.001
	6	0.98032319536991	0.885571755328864	0.999979611747038	< 0.001
Tibia	GL	0.999174563007128	0.997542371647678	0.999993885319161	<0.001
	SD	0.987544310157914	0.943820723804021	0.999680679056981	< 0.001
	SD1	0.885136910168308	0.611162058288325	0.996781122276528	< 0.001
	Bd	0.732035887952909	0.323867422067144	0.991271674830876	< 0.001
	Dda	0.907581769172941	0.669749780608486	0.997453235297581	< 0.001
	Ddb	0.99430946785968	0.974495870501259	0.999855035353237	< 0.001
Astragalus	GLl	0.999867808368355	0.999385224699985	0.99999664766876	< 0.001
	GLm	0.999176439494982	0.996240370364935	0.999979104797385	< 0.001
	H	0.949101096959984	0.791619955865441	0.998654754230955	0.0406
	DI	0.990913800487309	0.959498176810284	0.99976789435452	≥ 0.05 (0.0861)
	Bd	0.997911134229924	0.990274367566531	0.999946942065534	< 0.001
Calcaneum	c	0.959137534769449	0.774241132161043	0.999956648806104	0.0027
	d	0.922153672939921	0.622656620526543	0.999914661673294	< 0.001
	B	0.922890478439691	0.629005782899757	0.999914823266605	0.0067
	Dd	0.996416082719134	0.976892939308255	0.999996318834249	< 0.001
	Bd	0.924714429660189	0.72108735222619	0.997962212101553	≥ 0.05 (0.2047)
Metatarsus	GL	0.956915481062821	0.753641418774098	0.999953865205011	< 0.001
	SD	0.99547731432205	0.979651694954447	0.999884901063488	≥ 0.05 (0.2414)
	SD1	0.961839819428548	0.842834272321444	0.999001290366649	≥ 0.05 (0.1201)
	Bp	0.977086831379334	0.901496169523222	0.999407960946967	0.0014
	Dp	0.863669423261341	0.543713861510176	0.996054319121634	< 0.001
	BFd	0.982803348814844	0.918275744714498	0.999556836633846	< 0.001
	a	0.976250213801474	0.899034904400269	0.999385648798088	0.0043
	b	0.978883041128991	0.907668205347719	0.99945451866827	< 0.001
	l	0.997736019082914	0.988967428285356	0.999942472001034	≥ 0.05 (0.1242)

	2	0.999066308791611	0.995568674958832	0.999976301716891	0.0186
	3	0.997981865814919	0.990854278977674	0.999948746872654	< 0.001
	4	0.976503163162572	0.89954280957164	0.99939206144003	< 0.001
	5	0.997355448446153	0.986911961528576	0.999932752654252	0.0285
	6	0.996143974399747	0.981327889103473	0.999901852788924	< 0.001

Table 1. Inter-observer reliability estimates for all measurements across different anatomical elements from the three analysed specimens.

Element	Measurement	ICC	Lower CI	Upper CI	<i>p</i> value
Scapula	SLC	0.998017492252115	0.9969603571049	0.999918429014646	< 0.001
	GLP	0.999082284000876	0.998081529862206	0.999806090203964	< 0.001
	LG	0.998124403454768	0.996071226517356	0.999794020503819	< 0.001
	BG	0.998607528177046	0.997691903529594	0.999832598813459	< 0.001
	ASG	0.998849594941068	0.997819354877466	0.999645749414703	< 0.001
Humerus	GLC	0.998922088365431	0.997327779203704	0.999982587172965	< 0.001
	SD	0.997000718726439	0.996195005945303	0.998268530348392	< 0.001
	SD1	0.997374837482067	0.99395008444361	0.999804775916704	< 0.001
	BT	0.998464983132391	0.997945635864406	0.998920067981739	< 0.001
	Bd	0.998601167013423	0.998365834541977	0.998773908409192	< 0.001
	HTC	0.996788728279817	0.994991370827506	0.999723757997254	< 0.001
	HX	0.997200122334161	0.996883372711337	0.997463105477592	< 0.001
	HTL	0.996187078171646	0.993444444444444	0.998195589218089	< 0.001
Metacarpus	GL	0.999276991851092	0.997955025848451	0.999977827255846	< 0.001
	SD	0.995884994190517	0.994525693329696	0.998572023561611	< 0.001
	SD1	0.99745236480629	0.997232322583899	0.99764010689149	< 0.001
	Bp	0.998753466206853	0.9975488	0.999716763022218	< 0.001
	Dp	0.998520153403192	0.997183044335906	0.999591258848877	< 0.001
	BFd	0.998177537146729	0.997388058214908	0.998581326160824	< 0.001
	a	0.997524010659649	0.996325045866498	0.9987548828125	< 0.001

	b	0.997623215006023	0.994403551160308	0.999329370522786	< 0.001
	1	0.998148690846848	0.996781647729314	0.999151810635299	< 0.001
	2	0.997298649755776	0.99611666233584	0.999476513342255	< 0.001
	3	0.997714245072542	0.996805515656421	0.998467356934714	< 0.001
	4	0.997891953192544	0.99568345323741	0.999020759013514	< 0.001
	5	0.997360770213646	0.995560513285557	0.999455574948052	< 0.001
	6	0.997978270087526	0.99688501245995	0.999673008323425	< 0.001
Tibia	GL	0.999174563007128	0.997542371647678	0.999993885319161	< 0.001
	SD	0.99761654379144	0.995831656910059	0.999621928166352	< 0.001
	SD1	0.996774404030966	0.996164048153438	0.99716630880247	< 0.001
	Bd	0.997819790633323	0.997234358812666	0.998970797941596	< 0.001
	Dda	0.998518820839595	0.997908637738496	0.999420756861476	< 0.001
	Ddb	0.997865131366986	0.995522211720227	0.999206611570248	< 0.001
Astragalus	GLl	0.997329367948727	0.993574744420633	0.999971920839726	< 0.001
	GLm	0.996284710608565	0.993739285151653	0.999611393097309	< 0.001
	H	0.996638147441172	0.995498787195376	0.998308564332639	< 0.001
	DI	0.99828440511702	0.997342903998334	0.99959342339057	< 0.001
	Bd	0.998424687466066	0.997517738327723	0.999780091628488	< 0.001
Calcaneum	c	0.996376095905786	0.99535932278904	0.997346389304863	< 0.001
	d	0.997623295138574	0.996605555468762	0.999397474277308	< 0.001
	B	0.994529408063123	0.988679115196783	0.998190660239122	< 0.001
	Dd	0.997442164878804	0.995538501560874	0.998700348056048	< 0.001
	Bd	0.997419644097481	0.996762105524211	0.997764925007352	< 0.001
Metatarsus	GL	0.999302500325106	0.997923591812076	0.999994547442251	< 0.001
	SD	0.998376011868769	0.997066326530612	0.999255464814905	< 0.001
	SD1	0.99830939514967	0.9976802171443	0.999200951248514	< 0.001
	Bp	0.99817141220523	0.997172225985516	0.999895096755563	< 0.001
	Dp	0.997148124954235	0.995661993891456	0.999822537710736	< 0.001
	BFd	0.998530190016876	0.997584361805334	0.999151960345188	< 0.001
	a	0.998047794095296	0.996421158773872	0.999035168089153	< 0.001

	b	0.997805425760575	0.995438660887104	0.99932854442344	< 0.001
	1	0.997025642192717	0.995719620798621	0.998357954568457	< 0.001
	2	0.998414541660454	0.997432627097924	0.998925078202607	< 0.001
	3	0.998176745126049	0.996019169134031	0.999393990404848	< 0.001
	4	0.997937447042115	0.995619682564267	0.999640236686391	< 0.001
	5	0.996964922900909	0.99583157227388	0.997947485207101	< 0.001
	6	0.995922752287904	0.995111715724208	0.997297948387668	< 0.001

Table 2. Intra-observer reliability estimates for all measurements across different anatomical elements for the three analysed specimens.