prosthesis

MDPI

Article

A Real Time Delphi Study on the Challenges and Adverse Events to Percutaneous Osseointegrated Implant Integration and Long-Term Fixation in Limb Amputation

Kirstin Ahmed [1,*] and Max Ortiz-Catalan [1,2,3]

[1] Center for Bionics and Pain Research, Mölndals sjukhus, 43130 Mölndals, Sweden
[2] Prometei Pain Rehabilitation Center, 21000 Vinnytsia, Ukraine
[3] Center for Complex Endoprosthesis, Osseointegration, and Bionics, 04136 Kyiv, Ukraine
* Correspondence: kirstinahmed@gmail.com

Abstract: Percutaneous Osseointegrated Implants (POI) affix artificial limbs to the body after amputation. Several adverse events challenge mainstream uptake of this technology. This study aims to achieve a consensus regarding "*the challenges and adverse events to POI integration and long-term fixation in limb amputation*". We sought a panel of clinical experts divided by profession into surgical, clinical, or clinical academic categories. We used a real time eDelphi method to develop consensus on both the challenges and adverse event items, enabling anonymity, iteration, controlled feedback, and statistical aggregation of group responses. The full panel agreed that the most impactful items are amongst 10 key challenges and eight adverse events. Panellists were in consensus regarding the five most impactful challenges, which were, in decreasing order: *patient selection*, *absence of a multidisciplinary team*, *design of the implant*, *soft tissue stability* and an *experienced surgical team*. Panellists considered the five most impactful adverse events, in decreasing order, to be the following: *no biological fixation*, *deep infection*, *aseptic loosening*, *no mechanical fixation*, and *implant breakage*. Consensus was obtained on *implant breakage* and *deep infection* items. The proportion of consensus from the whole panel across all items was in line with the literature, and we observed an improvement in consensus once the panel was stratified based on job, expertise and implant system.

Keywords: percutaneous osseointegrated implant; bone anchored implant; direct skeletal fixation; amputation prosthesis; osseous integration; osseointegration; serious adverse events; orthopaedic Delphi study; real-time Delphi process; implant fixation

Citation: Ahmed, K.; Ortiz-Catalan, M. A Real Time Delphi Study on the Challenges and Adverse Events to Percutaneous Osseointegrated Implant Integration and Long-Term Fixation in Limb Amputation. *Prosthesis* **2024**, *6*, 1262–1280. https://doi.org/10.3390/prosthesis6050091

Academic Editor: Marco Cicciu

Received: 2 August 2024
Revised: 15 September 2024
Accepted: 24 September 2024
Published: 21 October 2024

1. Introduction

Percutaneous Osseointegrated Implants (POI) are an alternative to prosthetic sockets for the attachment of artificial limbs to the body after amputation. Forty years ago, Von Recum described five principle failure modes of POI: marsupialization (epithelial downgrowth), premigration, mechanical avulsion (mechanical induced failure), infections and abscess formation [1]. Despite intense research, the current literature still documents similar adverse events. For example, with respect to marsupialization, early research focused on achieving a dermal seal [2] akin to a deer antler structure connecting the pedicle to adjacent tissues. The absence of such a dermal seal acting as a microbial barrier has been associated with infections and hardware failure [3]. However, recent research suggests that improved surgical and rehabilitation techniques, particularly stabilization of the chronic wound environment at the percutaneous skin interface, have minimized this and other adverse events [4]. Reviews of the literature recommend better clinical management of low-grade infection and improved implant designs relative to loading regimens [5–13]. These recommendations rest on a preponderance of literature at evidence level $\leq$ III [14] that include heterogeneous study design, variable follow-up times and reported metrics/outcomes, and unequal research activity/output across the field. The

lack of consensus highlights the need to universally align and qualify recommendations to advance our collective knowledge, which then can be used to establish standards to address factors impacting osseous integration (mechanical and biological) and long-term fixation (referring to implant longevity in the host bone). Achieving a consensus regarding the challenges and adverse events associated with POI could potentially guide stakeholders to develop solutions to the obstacles preventing the widespread adoption of POI. This requires a universal cause-and-effect perspective; for example, a prosthetic component breakage without a holistic look at failure modes, such as asymmetric loading, would likely result in an incomplete picture. Similarly, outcomes limited to a single country or implant system do not represent an assessment of POI as a treatment or technology. For all these reasons, it has been difficult to conduct a meta-analysis, and thus a cause-and-effect consensus must be reached via alternate methods.

The purpose of this study is to develop a consensus on "The challenges and adverse events to POI integration and long-term fixation in limb amputation". *Consensus* is not a voting system, it is a method by which all actors can freely discuss opinions, and as a means for structuring group discussion and raising issues for debate, thus does not mean that a correct opinion has been found [15,16]. The Delphi method characteristics are anonymity, iteration, controlled feedback, and statistical aggregation of group responses [17]. A variation on the traditional Delphi method is an electronic (e) roundless 'real time' Delphi method that facilitates direct interaction between panellists and immediate feedback whilst generating equally robust results compared to traditional methods [18]. The method is well suited to this application since it can be used where evidence is ambiguous or lacking, enables a structured panel, and can be flexible in terms of geographical location; thus, a large scale global study, such as this one, is feasible [19] as well as being time- and cost-effective. The findings will potentially accelerate research and technology breakthrough(s) in the fields of POI design, surgical technique, patient selection, and rehabilitation protocols.

When considering osseous implant *integration*, we refer to immediate mechanical fixation at implantation (primary, 1°) and biological integration, i.e., osseointegration (secondary, 2°) fixation, see Figure 1. *Long-term fixation* is predominantly influenced by bone remodelling and stabilisation, which in turn are regulated directly by mechanical loading and indirectly by bone condition and metabolic health. Given the interdisciplinary nature of POI design, surgery, and rehabilitation, an expert panel crossing multiple knowledge domains is essential to develop standards to optimise osseous integration and long-term fixation. The panel should be able to draw on their experience and data from POI clinical trials, patient contact, long-term follow-ups, rehabilitation, and device- or procedure-related complications to reach a consensus on the causes and effects of POI failure.

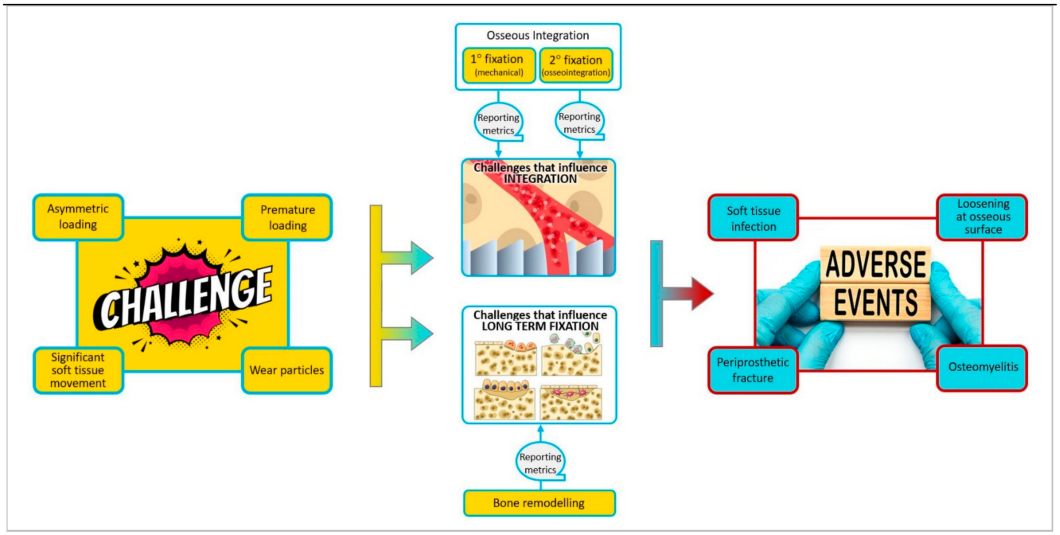

Figure 1. Illustrating the cause (challenges) and effect (adverse events) model with examples of each.

Note that there may be multiple adverse events resulting from an individual challenge, many challenges leading to an individual adverse event, or a combination of challenges resulting in many adverse events.

2. Materials and Methods

2.1. Preparation

2.1.1. Software Platform

The commercial software eDelphi (Metodix Ltd., Helsinki, Finland) was selected as the most suitable based on an analysis of electronic Delphi software tools [20].

2.1.2. Piloting

Prior to launching the eDelphi study, our research group tested the software platform, and two multidisciplinary focus groups reviewed information documents (focus groups included researchers and clinicians from different countries).

2.1.3. Expert Panel Inclusion Criteria

We sought a panel of clinical experts divided by profession into either a surgical, clinical, or clinical academic category. Within each category, panellists were asked to self-certify that they met the stipulated inclusion criteria (Table 1). The study was divided into two phases, an *exploratory* and an *evaluation* phase as is recommended for large public health eDelphi studies [21], see Figure 2.

Table 1. Panellist inclusion criteria and experience thresholds for all expert categories.

	Specialities	**Inclusion Criteria**	**Expertise**	
			Required	**Substantial**
SURGEON	Design, Plastic, Rehabilitation, Trauma, Orthopaedic, Oncology, Limb Salvage/ Reconstruction, Infection, other.	Experts must be medically and surgically certified in their own country and be a currently practicing surgeon.	Experts must have been the primary surgeon for between 3 $\leq$ 10 POI implants in patients treated for limb amputation or an assisting surgeon for at least 5.	Experts must have been the primary surgeon for > 10 POI implants in patients treated for limb amputation.
CLINICAL	Physiatrist, Physiotherapist, Certified Prosthetist, Surgical Nurse, Rehabilitation Therapist, Pain Therapist, other.	Experts must have direct patient contact as their main occupation and specialise in the rehabilitation and/or follow-up treatment for POI in patients treated for limb amputation. If other, they must have had equivalent clinical training.	Experts must be involved in the rehabilitation and/or follow-up treatment of between 5 $\leq$ 15 patients fitted with POI for limb loss for at least six months per patient.	Experts must be involved in the rehabilitation and/or follow-up treatment for > 15 patients fitted with POI for limb loss for at least six months per patient.
CLINICAL ACADEMIC	Professors, Directors, other.	Experts must at least be medically qualified in their own country but not necessarily currently practicing. They should be the PI, or a senior member in a group, running a registered clinical trial/investigation associated with POI in patients treated for limb amputation.	Experts must have run between 1 $\leq$ 3 registered clinical trials/ investigations associated with POI in participants with limb amputation.	Experts must have run> 3 registered clinical trials/investigationsassociated with POI in participants with limb amputation.

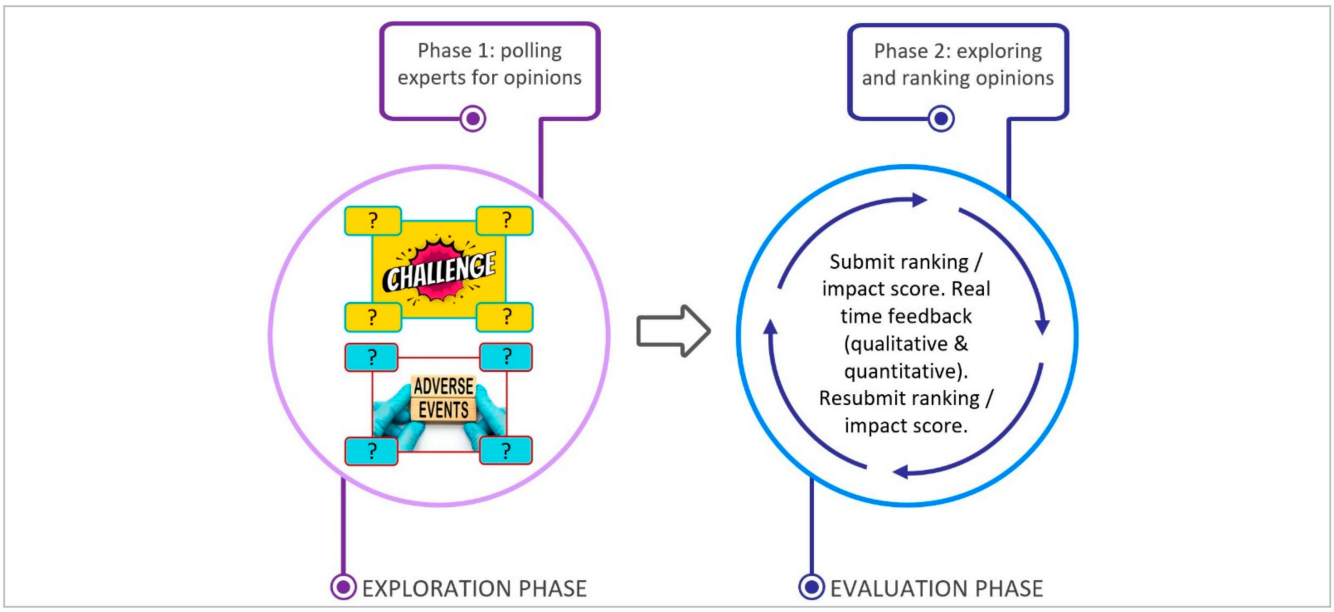

Figure 2. The two-phase real-time eDelphi process used in this study.

2.2. Phase One (Exploratory Phase)

2.2.1. Expert Panel Recruitment

Stage one: We performed a literature search in MEDLINE and Google Scholar from 2013 to 2023 using the following search terms: (osseointegrat* OR boneanchored OR bone anchored).ti. AND (arm OR upperlimb* OR upper extremit* OR transhum* OR transrad* OR leg OR lower limb* OR lower extremit* OR transfem* OR transtib*). Titles were screened, and studies that only included cadavers or were otherwise in vitro were excluded. All authors of the remaining studies were cross-referenced against our inclusion criteria (Table 1); if they did *not* meet the inclusion criteria, they were removed, but if inclusion was not possible to ascertain or they passed, then they remained on the list. We performed an online search for these authors' email addresses. A similar screening was carried out on public clinical investigation or trial databases, conference, symposium and scientific meeting speaker and panellist email address lists. Finally, we incorporated email addresses from our collective professional networks. After removal of duplicates, an initial list of 275 worldwide experts were placed on our eDelphi expert panel invitation list.

Stage two: We snowball sampled the panel for referrals of experts from their networks. We cross-referenced these against our inclusion criteria and our initial panel list to remove duplicates. This yielded a further 16 invitees, whom we added to the final panel invitation list (291 total).

2.2.2. Invitation

We sent personalised invitations via the software platform over the initial 14 days of the eDelphi study. The initial panel list of 275 were invited on the first of these days; the remaining 16 invitations were sent every three days thereafter, as they were received from the snowball sampling. The invitation provided information on how the study timings would work and what was expected from participation and was summarised with an infographic. The invitation stated that handling of personally identifiable information was fully compliant with the EU General Data Protection Regulation 2016/679 (GDPR) and US Privacy Law. Invitees were able to actively decline the invitation.

2.2.3. Self-Certification

If invitees chose to continue, they were asked to self-certify via the software platform confirming they met the inclusion criteria and were subsequently accepted as verified

panellists. If they did not meet the inclusion criteria, they were asked to certify that they would exit the process via the software platform.

2.2.4. Expertise and Experience

Using definitions provided for degree of expertise (Table 1), we asked the panellists to state their level of expertise with any of 12 POI implant types (Table 2). If a panellist met expertise thresholds in both upper and lower limbs, they were asked to select only upper OR lower limb to consistently answer in the eDelphi process. If they wanted to answer for both, they were offered an additional log in.

Table 2. POI systems.

POI Systems
1. OPRA SYSTEM (lower limb)
2. OPRA SYSTEM (upper limb)
3. OPL (TYPE A) SYSTEM
4. OPL (TYPE B) SYSTEM
5. ILP SYSTEM
6. EEP/EEFP SYSTEM
7. POP SYSTEM
8. ITAP SYSTEM
9. COMPRESS SYSTEM
10. BADEL X OFI-C
11. BADEL X OFI-Y
12. BADEL X OTI
13. OTHER SYSTEM (lower limb)
14. OTHER SYSTEM (upper limb)

2.2.5. Populating the Research Question (Providing Opinions)

In the final part of this exploratory phase, we asked panellists to provide opinions on the challenges and adverse events to POI integration and long-term fixation in limb amputation. Panellists were asked to provide opinions that they felt were significant to the research question in any order. Panellists were reminded that participation was asynchronous and that they could return an unlimited number of times to update their responses until the end of the exploratory phase. Panellist opinion submissions in the exploration phase were only seen by the eDelphi study managers.

2.2.6. Terminology

We provided study definitions for the terms "challenges", "adverse events", "integration", and "long-term fixation" and recommended a list of consistent terminology to use to account for any differences in languages and interpretation (see Tables S1 and S2 of the Supplemental File). During the exploration phase, we noticed substantial variability in panellists' classification of adverse events, probably due to the lack of standardisation for adverse events and consistency in outcome reporting in orthopaedics [22]. We therefore included all submissions that fell into most classification systems reported in the literature. These were the Orthopaedic Surgical Adverse Events Severity System (OrthoSAVES) [23], the Clavien-Dindo Classification of Surgical Complications [24] and the complications of total knee and hip arthroplasty standardised list, definitions, and stratification developed by the Knee and Hip societies, respectively [22,25]. Additionally, we included adverse events related to the use of an investigational medical device (the Adverse Device Effect, ADE) and the US Food and Drug Administration (FDA) system; Manufacturer and User facility Device Experience (MAUDE) [26] database for malfunctions.

2.2.7. Management of the Panel

Experts who accepted the eDelphi invitation were sent up to three reminders during the exploration phase to partake in the study if a response had not been submitted.

2.2.8. Opinion Data Gathering and Processing

Between phases one and two, we conducted a simple thematic analysis by manually searching, reviewing, defining, and naming themes [27]. Thereafter, opinion data were imported into phase two if they had been submitted by four or more panellists and became an "item". Opinion data that were submitted by fewer than four panellists were discounted.

2.3. Phase Two (Evaluation Phase)

2.3.1. Consensus and Stability Thresholds

Delphi studies typically use a measure of central tendency and dispersion to determine if agreement exists [28]. The Inter Quartile Range (IQR), accounting for the middle 50% of observations, compensates for outliers and is commonly used in Delphi studies [29,30]. In our study, consensus was reached when $IQR_{stratified} \leq 1.5$ in the stratified results and $IQR_{all} \leq 2.0$ across the full panel of experts (accounting for the higher expected variation in responses between different clinical expert categories). This threshold is considered a suitable consensus indicator for similarly constructed studies [31,32]. Stability between traditional Delphi rounds is a secondary measure of consensus [33] and considered reliable since it does not look for resistance to natural centralisation of views [34,35]. One could consider stability a quasi-static measure of consistency in answers between successive rounds of the study [36,37], albeit a very undefined aspect of consensus [38]. It is sometimes used in the consolidated statistics that are fed back to panellists between rounds and/or as a stopping criterion in combination with a consensus measure. Stability in a real-time Delphi process is evaluated at the end [18,33] and so cannot provide information as to how we got there and is therefore somewhat different from a traditional Delphi process. Several stability measures have been used in the literature [28,33,35,39–42]; we used the coefficient of quartile variation, CQV (Equation (1)) which measures the relative dispersion based on interquartile range [18,33]:

$$CQV = \frac{Q3 - Q1}{Q3 + Q1} \quad (1)$$

where Q3 and Q1 refer to quartiles 3 and 1, respectively. Stability of the evaluation phase was reached when $CQV_{all} < 30\%$ or $CQV_{stratified} < 15\%$, a strategy similar to that of Makovec, Goetzinger, Ribaut, Barnestein-Fonseca, Haupenthal, Herdeiro, Grant, Jácome, Roque, Smits, Tadic and Dima [42]. CQV is commonly used with data that have been organised into intervals. We assume that the intervals are equal on our Likert scale (between the 5 subdivisions from "no impact" to "most impact" axes labels) and the data can be treated as ordinal. We used IBM SPSS Statistics for Windows, version 29 (IBM Corp., Armonk, NY, USA) for the statistical analysis and Microsoft® Excel® for the consolidated statistics. In addition to a quantitative stability measure, the real-time Delphi process can provide insight into how the group moves towards consensus (or not) via qualitative data in the form of a panellist discussion.

2.3.2. Panellist Ranking and Denoting Impact

Panellists were asked to perform 4 tasks (2 ranking and 2 impact tasks):

1. Ranking task 1: To rank the *challenge* items in terms of their deleterious effects to successful integration and long-term fixation.
2. Ranking task 2: To rank the *adverse event* items in terms of their deleterious effects to successful integration and long-term fixation.
3. Impact task 1: To individually denote the deleterious impact to successful integration and long-term fixation that each *challenge* item had on a unipolar five-point Likert scale (from 0, which was labelled "no impact", to 5, which was labelled "most impact").

4. Impact task 2: To individually denote the deleterious impact to successful integration and long-term fixation that each *adverse effect* item had on a unipolar five-point Likert scale.

Tasks 1 and 2 were presented on a page each. Tasks 3 and 4 were presented across several (37) pages that were formatted equally with one challenge or adverse item per page and a slider to move a marker up or down a 5-point Likert scale. A discussion box on each page was available, and all comments and markers (denoting impact score) were visible to all panellists in real time.

2.3.3. Management of the Panel

Experts who had accepted the invitation and had not submitted a response were reminded up to three times during the evaluation phase that the eDelphi process was live and participation was asynchronous and that they could return an unlimited number of times to update their ranking/impact score. Also, during the evaluation phase, two emails to the full panel were sent with a reminder of the threshold for consensus and that to achieve this required continual reassessment of their ranking/impact score (by logging in and updating). All impact scores were updated in real time and remained visible to all panellists in real time.

3. Results

3.1. Panellist Engagement and Processing

A total of 74 panellists (25.4% of those invited) accepted our invitation to participate. We removed 14 who did not supply any answers to either phase of the study, and six who self-classified as not meeting the inclusion criteria. The data from these 54 panellists were accepted as submissions of opinions (phase one) to the research question. Thereafter, a further 23 were removed from analysis as they did not enter (n = 20) or complete (n = 3) the impact scoring part of phase 2. Leaving a final group of 31 panellists in phase 2 (10.7% of those invited) whose data we have analysed in this study.

3.1.1. Phase One (Exploration)

Of the 54 panellists in phase 1, four submitted opinions from the perspective of upper limb amputation, 50 from lower limb amputation, 52 from experience in humans and two with experience in animals. After processing the data, as previously outlined, this resulted in 37 items made up of 18 challenges and 19 adverse events, listed alphabetically in Table 3:

Table 3. Items for the Delphi process on the challenges and adverse events to successful integration and long-term fixation of POI in amputation.

	Challenge	Code		Adverse Event	Code
1:	Absence of Multidisciplinary (MD) Teams	AMT	*1:*	Aseptic Loosening	AL
2:	Access (geographical) to expert Rehabilitation	AR	*2:*	Enthesopathy	E
3:	Bone Length	BL	*3:*	excess Soft Tissue Redundancy	STR
4:	Comorbidities	C	*4:*	Failure to Biologically fix (osseointegration)	FB
5:	Design of Implant	DI	*5:*	Failure to Mechanically fix in bone	FM
6:	EXpense of hardware/surgery/rehabilitation	EX	*6:*	Fall Injuries	FI
7:	Experience of Surgical Team	EST	*7:*	Implant Breakage	IB
8:	High BMI	BMI	*8:*	Infection (grade 1 + 2)	I12
9:	High level of Activity	HA	*9:*	Infection (grade 3 + 4)	I34
10:	Non-Compliant patient	NC	*10:*	Neuroma pain	N
11:	Osteopenic/osteoporotic bone	O	*11:*	Component Problems (percutaneous)	CP
12:	overall Patient Selection	PS	*12:*	Periprosthetic bone Resorption	PR
13:	Prosthetic Alignment	PA	*13:*	Periprosthetic Fracture	F

Table 3. *Cont.*

	Challenge	Code		Adverse Event	Code
14:	Rehabilitation & support quality/experience/reliability	R	*14:*	PLP	PLP
15:	Smoking	S	*15:*	Safety Connector problems	SC
16:	Soft Tissue Stability	STS	*16:*	Skin Healing at the stoma	SH
17:	SToma Care and regimes	STC	*17:*	SToma Drainage	STD
18:	Surgical Technique—skeletal	ST	*18:*	SToma Pain	STP
			19:	SToma Type (wet/dry)	STT

Panellists stated involvement in clinical trials from the Deutsches Register Klinischer Studien DRKS00011564 (Comparison of conventional socket attachment and osseointegrated prosthesis fixation for transfemoral amputees), DRKS00022412 (Skin movement around stoma of transcutaneous osseointegrated prosthesis systems during daily activities), and DRKS00031106. Also, NCT06134167 (A Study to Evaluate the Safety and Effectiveness Transdermal Compress Device in Participants with Transfemoral Amputations), NCT 02564432 (Microbiome and Innate Immunity with Percutaneous Osseointegrated Prostheses). Also, the FDA-approved Osseointegrated Prosthesis for the Rehabilitation of Amputees (OPRA) early feasibility studies at Transhumeral, Transfemoral and forthcoming Transtibial amputation levels (TAOS, TFAOS, and TOSS, respectively); and the Office for Veteran Affairs UK trials and NCT02491424 (Intraosseous Transcutaneous Amputation Prosthesis (ITAP) trial) [43].

3.1.2. Phase Two (Evaluation)

Of the final group of 31 panellists, three answered from the perspective of upper limb amputation, 28 from lower limb amputation, all from experience in human patients. Phase two was further split into a *ranking* and an *impact* part; we only report the impact part due to a poor user experience with the ranking part on the software platform (the in-software chart confused panellists, and/or panellists only partially ranked the lists of challenges or adverse events).

3.2. Impact, Consensus, and Stability Amongst All Panellists

Figure 3 displays challenge and adverse event item median impact scores and IQR_{all} and CQV_{all} for all panellists.

3.2.1. Consensus

Challenges that did not reach consensus across the whole panel ($IQR_{all} > 2.0$) were *smoking* and *expense* ($IQR_{all} = 2.2$ and 2.8, respectively, median = 2.5 and 2.0, respectively). *Smoking* divided the panellists, with some scoring a low impact, stating "*smoking seems to have a bigger impact on soft tissue problems than on successful integration and long-term bone fixation of POI*" and "*the medical risk of smoking is probably exaggerated in the Orthopaedic society worldwide...*" and "*. . .remains to be statistically demonstrated*". Whereas others felt that it "*negatively impacts vascularisation and bone healing. . .*" and is "*most important and one of the few factors in our cohort that has a negative impact on integration. Smoking is an absolute contraindication*" and that "*smoking has a major impact on both bone-healing and healing of the soft-tissues. . .*". While others left pragmatic commentary "*if you die from smoking illness before the implant has a chance to loosen then you've avoided loosening problems within the lifetime of the patient*". In terms of the *expense* item, some panellists took umbrage with the connection to the research question; "*. . .the expense of the implant is a barrier to access from a socioeconomic perspective. Hard to adjust this to a difficulty of the bone growing into/onto the implant*". Whereas others made the link; "*it is very important if a connector breaks, and then the patient cannot afford to purchase another. Then they have no fail safe or may do their own fix to keep walking*"

likewise; "*implants are very expensive, and this includes connector which is a restricting factor on long term management*".

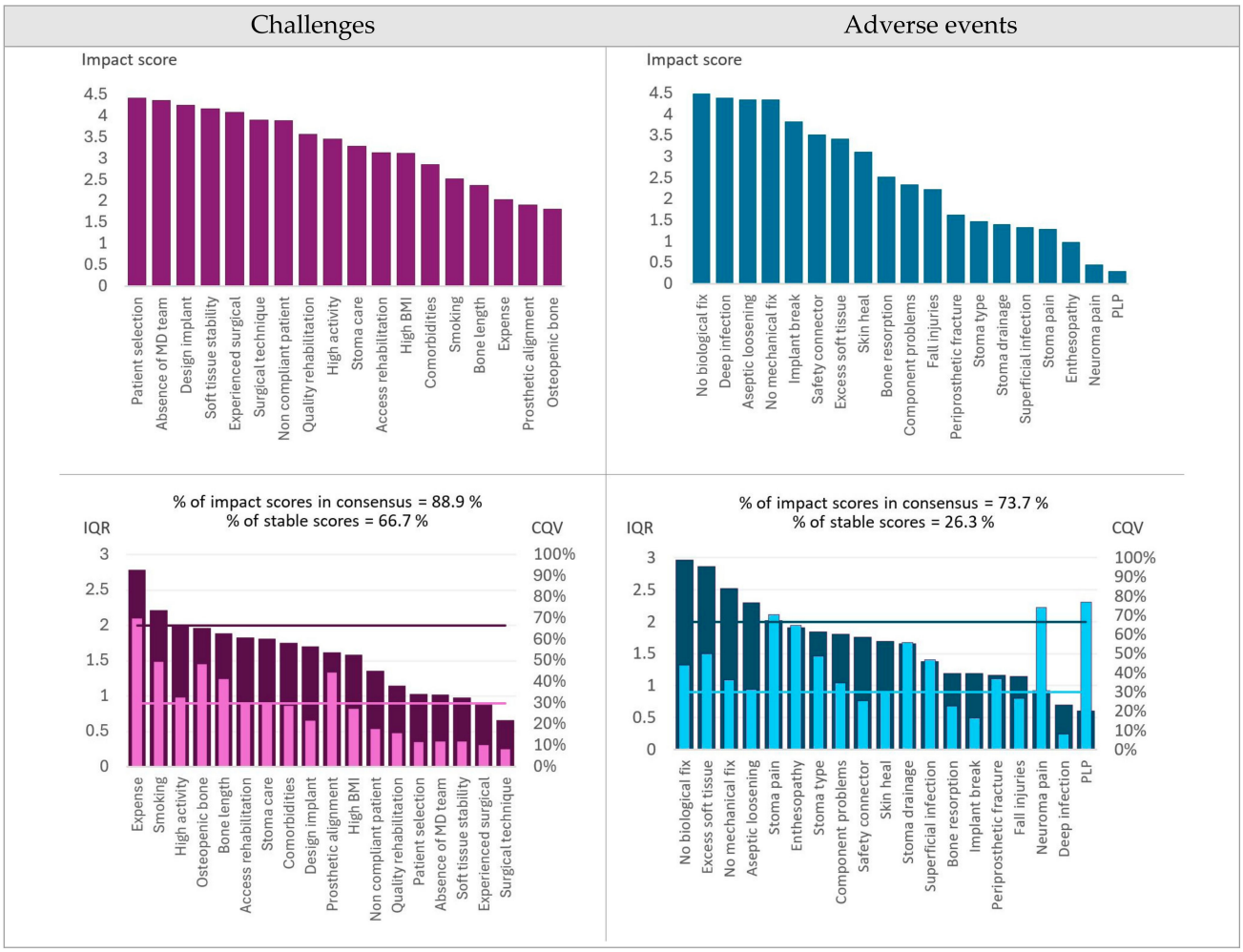

Figure 3. Top row: Median impact scores for 18 challenge items (pink) and 19 adverse events (blue), respectively. Bottom row: Dark pink and dark blue show IQR_{all} amongst all panellists. Light pink and light blue bars show CQV_{all} amongst all panellists. Solid horizontal lines at 2.0 IQR and 30% CQV indicate consensus and stability thresholds, respectively.

Adverse event items that did not reach consensus across the whole panel ($IQR_{all} > 2.0$) were *no biological fixation*, *excess soft tissue*, *no mechanical fixation*, and *aseptic loosening* (IQR_{all} = 3.0, 2.9, 2.5, and 2.3, respectively, median = 4.9, 3.4, 4.3, and 4.4, respectively). The *no biological fixation* item high IQR was the result of an almost binary split of scores at each end of the Likert scale. The *excess soft tissue* item similarly split the group with views such as "*this is huge*" and reasoning including "*it leads to instability at the stoma and subsequent problems*" to advice such as "*. . .trim more than you think. . .*" and commentary such as "*ignore the soft tissues at your peril, don't worry though, someone else will probably revise it*" while others thought that it was a "*source of complaint but not related to failure*". The *no mechanical fixation* item divided opinions again into low or high scoring groups with comments such as "*no primary fixation results in no/deficient osseointegration = loosening*" and "*. . .failure to achieve the primary purpose of the bone anchor*". The *aseptic loosening* item also resulted in difference amongst panellists with some not agreeing with the item's validity "*not sure if aseptic loosening even exists*" facing vehement opposition "*. . .it is important and if it happens you have fundamentally failed to achieve the whole point of osseointegration*" and "*aseptic loosening is an absolute disaster*".

3.2.2. Stability

Fourteen of the 18 challenge items were stable amongst all panellists (CQV_{all}); the other items (*expense, smoking, osteopenic bone*, and *prosthetic alignment*) were not. The *expense* item was unstable and did not achieve consensus amongst all panellists despite being supplied to phase 1 of the process by 16% (5) of the panellists. Despite all pain-related items (*stoma pain, enthesopathy, neuroma pain*, and *PLP*) achieving consensus on impact, CQV_{all} in these adverse events was approximately double the stability threshold. All panellists scored the *PLP* item ≥ 1.2 on the Likert scale, and views aligned in the discussion that nerve reconstruction was a solution.

3.3. Stratification of the Panel

3.3.1. Job & Expertise Stratification

Of the 31 panellists, 15 were surgeons (nine with *required* and six with *substantial* experience), five were clinicians (with *substantial* experience), two were clinical academics (with *required* experience), one panellist stated they met all three expertise categories at both experience levels, and eight were unclassified, i.e., did not state their expertise category but self-certified as meeting the inclusion criteria, totalling six sub stratifications.

3.3.2. Implant System Stratification

Of the 31 panellists, nine used the OPRA implant system, five the Osseointegrated Leg Prosthesis (OLP) system, seven were users of multiple systems (Mixed), eight did not classify (Unclassified), one used the ITAP system, and one used the Badel system, totalling six sub stratifications.

3.3.3. Highest Impact Scores

Figure 4 shows the highest impact scoring challenges and adverse event items across all panellists and amongst all sub stratifications in the stratifications. Note that the top five impact scoring challenges across all panellists come from about half (10) of the challenges (PS, STS, AMT, EST, ST, R, NC, DI, C, HA). Similarly, the top five impact scoring adverse events across all panellists come from about half (8) of the adverse events (I34, FM, FB, AL, IB, SC, STR, SH).

Challenges that were considered the most impactful across all panellists (row 1, Figure 4), in decreasing order, were *patient selection*, *absence md team*, *design implant*, *soft tissue stability* and *experienced surgical team*. Adverse events that were considered the most impactful across all panellists (row 1, Figure 4), in decreasing order, were *no biological fix*, *deep infection*, *aseptic loosening*, *no mechanical fix*, and *implant breakage*.

3.3.4. Impact, Consensus, and Stability within Stratifications

All stratified raw data and histograms of impact, consensus, and stability are in the Supplementary Materials (Tables S3 and S4, and Figures S1–S4, respectively). A summary of the percentage of items that are in consensus and stable amongst all sub stratifications compared to all panellists (non-stratified) is shown in Figure 5.

In terms of the 18 challenge items, we found the highest percentage of consensus and stability amongst the sub stratifications clinical *substantial* (100%, 72.2%, respectively) and academic *required* (94.4%, 88.9%, respectively). The least consensus (50%) was obtained in both the unclassified by job & expertise and the unclassified implant system sub stratifications. The lowest stability observed (all below 33%) amongst the challenge items was in the surgical *substantial*, OLP and mixed implant system sub stratifications.

In the adverse events, the pattern of consensus and stability amongst the stratifications was not dominated by one or two sub stratifications in the same way. All sub stratifications attained a consensus in over 68% of the 19 items except for the OPRA and mixed implant system sub stratifications. Stability was relatively low across all sub stratifications (as was the case across all panellists) except for the academic *required* sub stratification, and none attained stability in the adverse events above 50%

	Top 5 CHALLENGES					Top 5 ADVERSE EVENTS				
	1st	2nd	3rd	4th	5th	1st	2nd	3rd	4th	5th
All panellists	**PS**	**AMT**	**DI**	**STS**	**EST**	**FB**	**I34**	**AL**	**FM**	**IB**
Surgeon *required*	PS	DI	EST	ST	NC	FB	I34	AL	SC	FM
Surgeon *substantial*	STS	EST	AMT	R	NC	STR	I34	IB	SC	SH
Clinical *substantial*	NC	AMT	C	STS	PS	FB	FM	I34	AL	IB
Academic *required*	AMT	PS	R	HA	ST	AL	SC	IB	FM	I34
Unclassified expertise	DI	PS	ST	EST	STS	AL	FB	FM	I34	SC
OPRA system	PS	EST	DI	NC	AMT	FB	AL	I34	SC	FM
OLP system	PS	AMT	STS	R	EST	I34	FB	STR	FM	IB
Mixed system	AMT	ST	EST	STS	PS	I34	AL	IB	FM	FB
Unclassified system	DI	PS	ST	EST	STS	AL	FB	FM	I34	SC
ITAP system	NC	STS	C	AMT	R	FB	FM	I34	IB	STR
Badel system	AMT	R	PS	STS	ST	I34	AL	FB	FM	IB

Figure 4 above: Top five median impact scores amongst all stratifications for all challenges and adverse events. Right: Key to abbreviations and the items' occurrence within the top 5 challenges or adverse events amongst all stratifications.

Challenge	Occurrence
Patient selection (PS)	9
Soft tissue stability (STS)	8
Absence MD team (AMT)	8
Experienced surgical team (EST)	7
Surgical technique (ST)	6
Quality rehabilitation (R)	5
Non-compliant patient (NC)	5
Design implant (DI)	4
Comorbidities (C)	2
High activity (HA)	1
Adverse event	
Deep infection (I34)	11
No mechanical fix (FM)	10
No biological fix (FB)	9
Aseptic loosening (AL)	8
Implant breakage (IB)	7
Safety connector (SC)	6
Excess soft tissue (STR)	3
Skin heal (SH)	1

Figure 4. Above: Top five median impact scores amongst all stratifications for all challenges and adverse events. Right: Key to abbreviations and the items' occurrence within the top five challenges or adverse events amongst all stratifications.

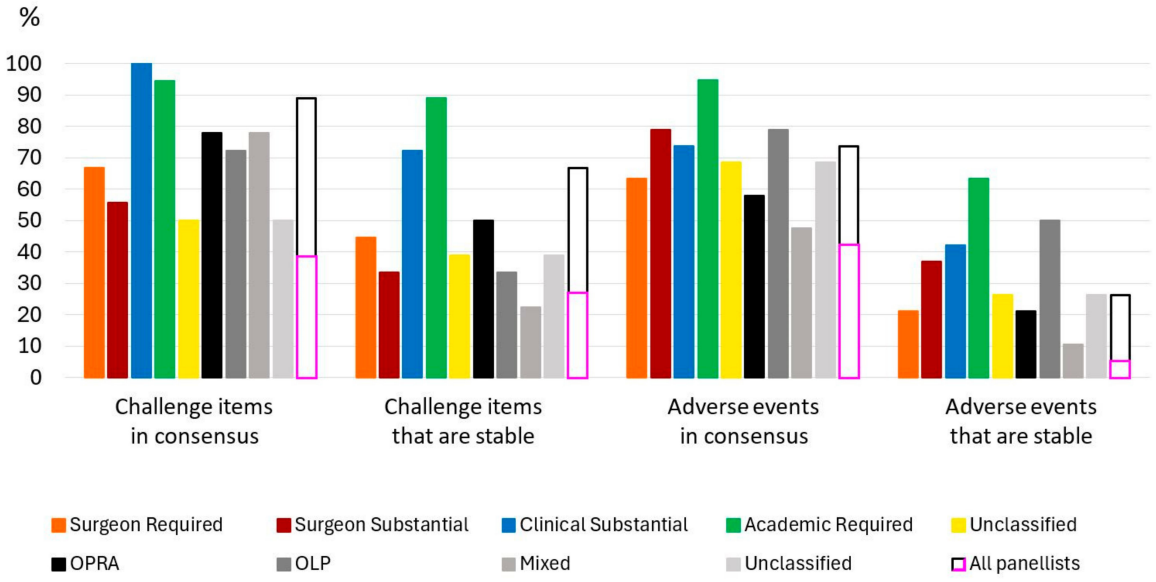

Figure 5. Percentage of IQR and CQV under thresholds for consensus and stability amongst all sub stratifications (thresholds were $IQR_{stratified} \leq 1.5$ and $CQV_{stratified} \leq 15\%$) and all (non-stratified) panellists (thresholds were $IQR_{all} \leq 2.0$ and $CQV_{all} \leq 30\%$ indicated with a black outline. Results using the stratified thresholds are indicated with a pink outline).

3.3.5. Statistical Analysis

We performed an independent samples Kruskal–Wallis one-way ANOVA two-sided test to determine the significance of the mean rank sum differences in impact scores between the stratified results. Note that we were unable to compare distribution of medians since the data were not distributed similarly between stratifications (from a visual inspection of box and whisker plots).

When stratified by job & expertise, we observed a statistically significant difference (at $\alpha = 0.05$, degrees of freedom = 4) in the mean rank sum differences in impact scores for the *implant design* and *surgical technique* items (p values = 0.026 and 0.043), see Figure 6a. Post hoc pairwise comparisons with the Dunn's test were performed. The *implant design* item (n = 29) revealed no significance between pairs after adjustment by the Bonferroni correction for multiple tests. The *surgical technique* item (n = 27) revealed pairwise significance between surgical *substantial* and unclassified sub stratifications (adjusted p value = 0.046). When stratified by implant systems, we observed a statistically significant difference (at $\alpha = 0.05$, degrees of freedom = 5) in the mean rank sum differences in impact scores for the *implant design* item (p value = 0.025), see Figure 6b. Post hoc pairwise comparisons with the Dunn's test were performed. The *implant design* item (n = 29) revealed pairwise significance between the mixed and unclassified sub stratifications after adjustment by the Bonferroni correction for multiple tests. The impact scores for mean rank sum differences for the remaining challenge items and all adverse events in the stratified results were not statistically significant.

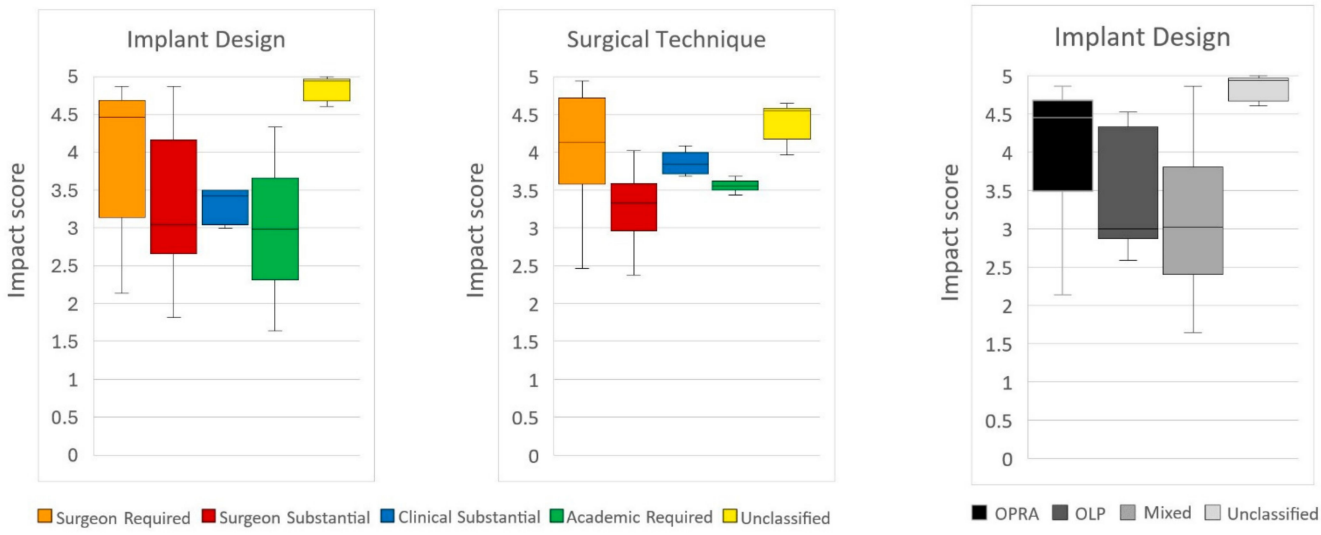

Figure 6. (**a**) Box and whisker plots of stratified (by job & expertise) impact scores where statistically significant difference in the mean rank sums was obtained; (**b**) Box and whisker plot of stratified (by implant system) impact scores where statistically significant difference in the mean rank sums was obtained.

4. Discussion

We sought to develop a worldwide consensus from clinical experts on the challenges and adverse events to POI integration and long-term fixation in limb amputation. It is accepted that panellists for these types of studies are challenging to retain and keep engaged [44,45]. We therefore opted for an expedited version of the traditional Delphi method to reduce the timeframe between exploratory and evaluation phases, as this version of the process has shown that panellists become more committed, thereby maximising the validity of results [46]. All versions of Delphi studies seeking consensus on clinical interventions will typically involve non-probability sampling techniques, which potentially

reduces representativeness. However, our panellist group was large and broad with a high-quality score as defined by Diamond, Grant, Feldman, Pencharz, Ling, Moore and Wales [47] (the use of this quality score usually requires the number of rounds stated, clear criteria for dropping items and stopping criteria other than rounds specified. These considerations are not applicable to real-time Delphi studies). We evaluated the stability of consensus by measuring the relative dispersion based on interquartile range between consecutive rounds (CQV). In a real-time Delphi study, this is done at the end of the consensus process. In addition to this quantitative measure of stability, we were able to observe how the group moved towards consensus (or not) via qualitative interim data in the form of a panellist discussion.

Consensus thresholds in Delphi studies varies widely, and measuring standards for reaching consensus do not yet exist [28,48]. Varndell, Fry and Elliott [33] state consensus was obtained if ≥83% of panellists ranked the item ≥ 7 on a 9-point Likert scale. Loughlin and Moore [49] and McKenna [50] defined consensus as 51% agreement among panellists, whereas Sumsion [51] recommends 70% and Green, Jones, Hughes and Williams [52] opted for 80%. Furthermore, reporting of consensus in studies is poor, with only 75% of 100 studies between 2000 and 2009 defining consensus in a recent review [47]. This definition was stated *a priori* in 89% of the studies; however, measuring standards for reaching consensus were not specified in 33% of the studies. In our study, we defined consensus *a priori* as $IQR_{stratified} \leq 1.0$ and communicated this to the panellists. However, post hoc, we noted that both references [31,32] upon which we aligned our consensus thresholds due to similarity in study design had used an interquartile *deviation*, IQD = 1.0 rather than IQR = 1.0 as their measure of consensus. Since IQD is half the value of IQR, this was in fact a more stringent requirement for consensus (IQD of 1.0 = IQR of 2.0). We therefore adjusted our thresholds for analysis in this study after data collection to $IQR_{all} \geq 2.0$ and $IQR_{stratified} \geq 1.5$, as outlined in the methods. Our consensus results across all panellists were within ranges in the literature (IQR_{all} = 88.9% of challenge items and 73.7% of adverse events). Once stratified, these results improved dramatically, as shown in Figure 5, where every item achieved a higher percentage of panellists in consensus when judged by the same threshold ($IQR_{stratified} \geq 2.0$) and a similar percentage achieved consensus when judged by a more stringent threshold ($IQR_{stratified} \geq 1.5$).

We intentionally designed for a heterogenous expert panel, as a wider range of perspectives leads to better performances and higher-quality responses in the Delphi process [53]. For example, in the panellist discussions, we observed differences in the way some items were interpreted, such as *prosthetic alignment* that was sometimes confused with implant stem misalignment. Without a thorough thematic analysis of the discussion, we are unable to see if different item interpretations are inter- or intra-stratified, and so panel heterogeneity is valuable. Of note is an unequal population size in each sub stratification, which is non ideal but hard to manage in the Delphi process, and moreover, had an impact on the statistical analysis. Items that are in consensus but that are unstable indicate a relatively large IQR compared to the sum of the quartile impact scores (considering Equation (1)). A relatively low sum of the quartile impact scores will result from a low Q1 and/or a low Q3. In the cases of the high instability amongst the pain-related items, the panellist discussion offers some insight. From zero and very low impact scores being the majority, with comments such as "*affects QOL but not OI*" and "*OI not likely to press on neuroma*", to less common and the highest scores "*. . .can be so disabling as to render the entire bone-anchor procedure a failure...*", there was a general agreement that nerve reconstruction in the form of TMR/RPNI helps; "*. . .TMR/RPNI has helped significantly*" and "*TMR/RPNI can address this in most patients*".

We noticed that there were some inconsistencies in the way panellists were scoring some items; this was particularly clear in panellist discussions on some adverse event items. For example, the *no biological fixation* item resulted in a bipolar spread of scores, making it almost impossible to achieve consensus. Very low scores, which were in the minority, were given because this item was "*rarely seen*" or "*not observed yet*", whereas the majority

gave the highest scores possible, stating "*a fundamental failure of the entire procedure*" and "*. . .impact would be huge*". Similarly, the no mechanical fixation item included panellists scoring low from an occurrence perspective, stating, "*never encountered so far"* and that it was "*uncommon*", with the misinterpretation spelled out by one panellist stating "*. . .i think we are inflating the impact of something that is exceedingly rare and that I have only seen once*". Whereas others scored highly from an impact perspective: "*. . .the sole purpose of OI surgery is lost"* and "*no primary fixation results in no/deficient osseointegration = loosening*".

Other examples of this were the *osteopenic bone* and the *enthesopathy* items. *Osteopenic bone* was considered of zero impact by some panellists, who cited "*. . .three published studies. . .*", to mid-range impact scores, including several comments relating occurrence: "*osteopenia is normal in amputees (non-acute). Severe osteopenia would concern me...*" and "*osteopenic bone is the norm for most amputees; so, integration can occur, it just requires more careful bone loading and longer healing time*" similarly "*osteoporosis not important so far in my practice except for pace of loading*" and "*. . .moderate osteopenia is standard in amputated limbs. No problem to OI*". Impact scoring inconsistencies, if reflected by the comments, were in the minority yet also spotted and commented on by one panellist: "*if the implant is loose, then there is no osseointegration; can't imagine any of the "experts" polled would agree that a loose implant wasn't very IMPACTFUL. . .*". In terms of the *enthesopathy* item, some panellists scored very low impact, stating it was "*. . .relatively common and can be quite annoying*" and "*affects QOL but not OI*", and others considered it to be more impactful, with comments such as "*. . .common and can be very disabling*" and "*impacts the entire OI experience due to nagging, activity related pain*" and capable of affecting the rehabilitation protocol, e.g., "*pain from enthesopathy can reduce the weight bearing*", whereas other mid-scoring panellists suggested *implant design* played a part "*far too common in F platform and less so in G but still happens. . .*".

There is a fine line between occurrence and impact; panellists were encouraged to draw on their expertise to denote impact, and the discrepancy in scoring perception was due to whether they considered personal experience/occurrence to be the sum total of their experience or not. We observed that panellists scored occurrence rather than impact to a greater degree in the adverse event items. One might question why all sub stratifications had the *no mechanical fixation* and *aseptic loosening* items in their top five except for the surgeon *substantial* sub stratification. A possible deduction is that a higher proportion of panellists in this sub stratification scored from an occurrence rather than from an impact perspective. However, we should note that there was no way to be sure of this since all comments were anonymised; suffice it to say that we have demonstrated the importance of capturing panellist discourse and highlighted one of the fundamental challenges with opinion-based polls such as the Delphi process.

We filtered out items received in phase 1 of the process that were not submitted by at least four panellists and subsequently removed any that did not fall into the categorizations we used for adverse events. However, this remained still open to interpretation, and there was some useful debate around terminologies and the included items, such as "*the word 'challenges' apparently is differently interpreted by the participants of this panel, which is understandable and reasonable. . .it is very hard to co-rank biological/technical challenges on the same list as socioeconomic challenges to initiating care".* In a similar vein, there were panellists who also felt we should have gone further in the processing (removal) of opinions between phases of the study: "*. . .absence of multi disciplinary teams and access to rehabilitation should not really be here as these are a minimum requirement of doing this surgery in my opinion. It is negligent to do it any other way*" and "*PLP and neuroma pain are indeed a severe problem for some patients, but do not interfere with "successful integration and long term bone fixation" which the question ask about. The implant can be well integrated, but the pain can remain very problematic for the patient*." Some of these comments, in particular the ones relating to pain, were reflected in the results, in so much as the top five impact scoring challenges amongst all panellists and all stratifications came from about half of the challenges, none of which included pain items. Likewise, the top five impact scoring adverse events across all panellists and all stratifications came from about half of the adverse events. Further useful feedback on the

Delphi process of scoring/ranking an item was challenged since in the clinic "*. . .there are issues we are confronted with due to our patient selection (Soft tissue issues might be interlinked with comorbidities and patient selection, smoking, BMI,* etc.)".

When contrasting the results from our study with the POI outcome literature, we should be mindful that due to the large-scale global nature of this Delphi process and the fact that the field is evolving so quickly, we should expect some differences. For example, to date, it has been reported that the occurrence of superficial infections has been linked to the issue of soft tissue redundancy [54,55]. We did not see soft tissue redundancy (*excess soft tissue*) in the top five adverse events in our Delphi study. Similarly, component part failure is often reported in the POI literature [55–57] and although still reported in very recent longitudinal retrospective data [58], does not feature in the top five adverse events in our Delphi study. On the other hand, we saw *no mechanical fix* and *deep infection* as the top two items in our Delphi study. This is in line with most of the current literature [11,59,60]. Our results show that the Delphi process has united opinion and globally demonstrated the impact of these items despite the literature reporting using heterogeneous study design, variable follow-up times and reported metrics/outcomes, and unequal research activity/output across the field.

When stratified by job & expertise, only the surgeon *required* sub stratification and the unclassified expertise sub stratification had *design implant* amongst their five highest impact scoring items. The same two sub stratifications did not have *absence MD team* in the five highest impact scoring items, unlike the rest of the sub stratifications in this stratification. The surgical sub stratification was the only one containing panellists from two levels of experience; and we observed some difference between their top five items (except for *non-compliant patient* that was in position five in both). The less-experienced surgeons (surgeon *required*) did not score *soft tissue stability* in their top five (it came seventh), whereas the more-experienced surgeons (surgeon *substantial*) had it in position one. Discussion on the software platform overall was left by those who considered it to be of utmost importance, such as "*this is one of the most important issues in bone-anchor surgery. Achieving it is like searching for the holy grail...*" and "*possibly the single most important aspect of the procedure, healing and patient outcomes*", including explanations such as "*achieving very low levels mobility of tissue at metal interface is most important to prevent significant complications, exudate, infections, connector issues with tissue*" and "*. . .movements pumps secretion and bacteria and results in infections*". *Patient selection* was considered the most impactful item amongst all panellists and occurred the most often when stratified. *Patient selection* was the item that the surgeon *required* sub stratification considered the most impactful, as did both the OPRA and OLP implant system sub stratifications. The clinical *substantial* sub stratification and the ITAP system sub stratification considered *non-compliant patient* to be the most impactful item, which, like *patient selection*, indicates how important the panellists feel that only the *right* patients are offered this treatment. Comments included "*the most important factor for compliance and good chance of healing, taking care of their limb and presenting early if having problems*" and "*patient selection is one of the most important factors for success in bone-anchor surgery*" for reasons such as "*it is very important to select patients who can/will follow instruction*" and obstacles such as "*. . .challenging to get all surgeons and physiatrists to agree on*" and "*non-compliant patients who failed to attend surgical and rehabilitation appointments did less well*".

All sub stratifications apart from surgeon *substantial* and academic *required*, had *no biological fix* in their five highest impact scoring items (surgeon *substantial* ranked this item 14th, and academic *required* ranked it 18th, which we must assume is an occurrence rather than an impact score). Similarly, all sub stratifications apart from surgeon *substantial* had *no mechanical fix* in their five highest impact scoring items (surgeon *substantial* ranked no mechanical fix in position 15). The *deep infection* adverse event appeared in all sub stratifications in the five highest impact scoring items and was in the first position for three of the implant system sub stratifications (OPL system, Mixed systems, and Badel system). Comments highlighted the severity of this item: *". . .usually results in removal of*

implant", and some referred to occurrence "*not common but huge impact on residual limb and loss of OI*" and "*. . .a deep infection which affects the entire implant resulting in loosening and loss of the osseointegration is a very rare event indeed*". *Aseptic loosening* featured in all sub stratifications in the job & expertise stratification top five highest impact scores apart from surgeon *substantial* (where it ranked ninth). The *implant breakage* adverse event item was the fifth-ranked adverse event amongst all panellists and featured in seven out of 11 sub stratifications' top five but never above position three. Comments suggested the severity and occurrence in equal measure: "*only come across one but impact was significant*" and "*. . .uncommon but it is serious and needs revision surgery*", along with several comments referring to the cause: "*. . .but risk is really with connectors*" and "*. . .suggests that someone was not being as careful as they could have been with use of the connectors to reduce wear on the implant*".

Overall, stratifying the panellists based on job & expertise and implant system has offered some valuable insight. It is somewhat expected that there were differences between the sub stratifications and particularly that items such as *implant design* and *surgical technique* were significantly different. We were somewhat limited by low populations in the sub stratifications (except for the surgeons) and we should interpret statistical results cautiously.

Reflecting on the final expert panel constellation, we consider that the thresholds for *required* and *substantial* expertise were appropriate. Although many invited panellists did not meet the inclusion criteria, those who did voiced a reliably expert opinion [16] and included several PIs in international clinical trials. Ideally, a larger group of experts would have participated, particularly in the clinical academics' stratification, of which there were only two, and none with *substantial* experience. Moreover, eight of the panellists were of *unclassified* expertise; in these cases, with larger and more specific stratifications, we would expect more robust results. In the stratification of the panel, once again, we recognise that improved study robustness would have been obtained with a larger panel. Although we were able to stratify into four POI systems, it was suboptimal that two of them (ITAP and BADEL) were only represented by one panellist each, and that we did not receive any panellists with the required experience using the EEP/EEFP, POP, or Compress systems. We would have liked to stratify by clinical trials; however, we did not receive enough data to investigate this. The Delphi process is designed to correct for the absence of conclusive data through the collective sharing of experiential knowledge of its expert panel [61], thus is not subject to the same validation standards as a scientific method. Instead, it should be viewed as a process that makes the best use of available information and relies upon experts being fully engaged with the process and motivated by the outcomes. It has some limitations, and this can extend to the way in which the panel is managed, including the degree of intervention on item selection considered reasonable. In a real-time Delphi process a dedicated software is required [18,20,42]; however, despite testing our software, we were not prepared for the large number of manager/panellist issues, glitches and challenges we faced in and during execution of the study. The next step towards clinical practice is to develop a similar set of functional and psychological outcomes, perhaps also employing a consensus method, to fill the picture on POI overall success.

5. Conclusions

Using the real-time Delphi method on a heterogenous group of clinical experts resulted in the top five challenge items considered the most impactful, in descending order, to be *patient selection, absence md team, design implant, soft tissue stability*, and *experienced surgical team*. Likewise, the top five adverse events considered most impactful, in descending order, were *no biological fix, deep infection, aseptic loosening, no mechanical fix,* and *implant break*. Although causality cannot be proved, we can say that these challenges contribute to these adverse events. More concretely, we can conclude that a heterogenous expert panel achieved consensus regarding the top five impact scoring challenge items and two of the five top-scoring adverse events (*deep infection* and *implant breakage*). We can suggest that the other three top-scoring adverse events (*no biological fix, aseptic loosening*, and *no mechanical*

fix) might not have achieved consensus only because of differences in scoring interpretations and that this was most prevalent in one sub stratification (surgeon *substantial*). We saw marked improvement in consensus and stability (with the same thresholds) once the panel was stratified, as one might expect; however, we would like to have had more representativeness in clinical and academic sub stratifications. Clinical Recommendation: We have demonstrated that despite panellists differing foci (surgical, rehabilitation, and/or clinical research) there is agreement that the most impactful challenges and adverse events reside in a narrow range of items (10 challenges and 8 adverse events). Therefore, we should focus on these cause-and-effect items to expedite research and technology developments, surgical technique, patient selection, and rehabilitation protocols.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/prosthesis6050091/s1, Figure S1: Impact scores (median) in all groups in both stratifications for all challenge items; Figure S2: Impact scores (median) in all groups in both stratifications for all adverse event items; Figures S3a–j: Bar charts of consensus (dark bar colour = $IQR_{stratifed}$) and stability (light bar colour = $CQV_{stratified}$) of the impact data stratified by expertiese & job indicating where the thresholds are (dashed line = $IQR_{stratified} \leq 1.5$ and solid line = $CQV_{stratified} \leq 15\%$ respectively); Figures S4a–h: Bar charts of consensus (solid bar colour = $IQR_{stratifed}$) and stability (patterned bar colour = $CQV_{stratified}$) of the impact data stratified by implant system indicating where the thresholds are (dashed line = $IQR_{stratified} \leq 1.5$ and solid line = $CQV_{stratified} \leq 15\%$ respectively). Table S1: Consistent terminology study definitions that were provided on the software platform to all panellists throughout the study [62]; Table S2: Study definition and intention notes that were provided on the software platform to all panellists throughout the study; Table S3: Impact scores for all challenges and adverse events from all panellists stratified according to job (and expertise); Table S4: Impact scores for all challenges and adverse events from all panellists stratified according to implant system.

Author Contributions: K.A. was responsible for Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Visualization, Writing—Original Draft Preparation, Writing—Review and Editing. M.O.-C. was responsible for Conceptualization, Funding Acquisition, Resources, Supervision, Writing—Review and Editing. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by the Promobilia Foundation (20500).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the manuscript, further inquiries can be directed to the corresponding author.

Conflicts of Interest: M.O.-C. has consulted for Integrum AB. K.A. declares no competing interests.

References

1. Von Recum, A.F. Applications and failure modes of percutaneous devices: A review. *J. Biomed. Mater. Res.* **1984**, *18*, 323–336. [CrossRef] [PubMed]
2. Pendegrass, C.J. Soft Tissue Attatchment to Orthopaedic Implants. Ph.D. Thesis, University of London, London, UK, 2005. Available online: http://search.proquest.com/docview/1428985547?accountid=14511 (accessed on 21 July 2024).
3. Tillander, J. Infections Associated with Percutaneous Osseointegrated Titanium Implants for Limb Prostheses. Ph.D. Thesis, University of Gothenburg, Gothenburg, Sweden, 2017.
4. Souza, J.M.; Mioton, L.M.; Harrington, C.J.; Potter, B.K.; Forsberg, J.A. Osseointegration of extremity prostheses: A primer for the plastic surgeon. *Plast. Reconstr. Surg.* **2020**, *146*, 1394–1403. [CrossRef] [PubMed]
5. Al Muderis, M.M.; Lu, W.Y.; Li, J.J.; Kaufman, K.; Orendurff, M.; Highsmith, M.J.; Lunseth, P.A.; Kahle, J.T. Clinically relevant outcome measures following limb osseointegration; systematic review of the literature. *J. Orthop. Trauma* **2018**, *32*, e64–e75. [CrossRef] [PubMed]
6. Atallah, R.; Leijendekkers, R.A.; Hoogeboom, T.J.; Frölke, J.P. Complications of bone-anchored prostheses for individuals with an extremity amputation: A systematic review. *PLoS ONE* **2018**, *13*, e0201821. [CrossRef] [PubMed]
7. Balzani, L.D.; Ciuffreda, M.; Vadalà, G.; Di Pino, G.; Papalia, R.; Denaro, V. Osseointegration for lower and upper-limb amputation a systematic review of clinical outcomes and complications. *J. Biol. Regul. Homeost. Agents* **2020**, *34*, 315–326.

8. Gerzina, C.; Potter, E.; Haleem, A.M.; Dabash, S. The future of the amputees with osseointegration: A systematic review of literature. *J. Clin. Orthop. Trauma* **2020**, *11*, S142–S148. [CrossRef]
9. Hoellwarth, J.S.; Tetsworth, K.; Rozbruch, S.R.; Handal, M.B.; Coughlan, A.; Al Muderis, M. Osseointegration for amputees: Current implants, techniques, and future directions. *JBJS Rev.* **2020**, *8*, e0043. [CrossRef]
10. Kunutsor, S.; Gillatt, D.; Blom, A. Systematic review of the safety and efficacy of osseointegration prosthesis after limb amputation. *J. Br. Surg.* **2018**, *105*, 1731–1741. [CrossRef]
11. Mohamed, J.; Reetz, D.; van de Meent, H.; Schreuder, H.; Frölke, J.P.; Leijendekkers, R. What are the risk factors for mechanical failure and loosening of a transfemoral osseointegrated implant system in patients with a lower-limb amputation? *Clin. Orthop. Relat. Res.* **2022**, *480*, 722–731. [CrossRef]
12. Overmann, A.L.; Aparicio, C.; Richards, J.T.; Mutreja, I.; Fischer, N.G.; Wade, S.M.; Potter, B.K.; Davis, T.A.; Bechtold, J.E.; Forsberg, J.A. Orthopaedic osseointegration: Implantology and future directions. *J. Orthop. Res.* **2020**, *38*, 1445–1454. [CrossRef]
13. Sreedharan, S.; Gray, S.; Bruscino-Raiola, F. Osseointegrated prostheses for lower limb amputees: A review of complications. *Australas. J. Plast. Surg.* **2021**, *4*, 56–62. [CrossRef]
14. Burns, P.B.; Rohrich, R.J.; Chung, K.C. The Levels of Evidence and Their Role in Evidence-Based Medicine. *Plast. Reconstr. Surg.* **2011**, *128*, 305–310. [CrossRef] [PubMed]
15. Tastle, W.J.; Wierman, M.J. Consensus and dissention: A measure of ordinal dispersion. *Int. J. Approx. Reason.* **2007**, *45*, 531–545. [CrossRef]
16. Hasson, F.; Keeney, S.; McKenna, H. Research guidelines for the Delphi survey technique. *J. Adv. Nurs.* **2000**, *32*, 1008–1015. [CrossRef] [PubMed]
17. Rowe, G.; Wright, G. The Delphi technique: Past, present, and future prospects—Introduction to the special issue. *Technol. Forecast. Soc. Chang.* **2011**, *78*, 1487–1490. [CrossRef]
18. Gnatzy, T.; Warth, J.; von der Gracht, H.; Darkow, I.-L. Validating an innovative real-time Delphi approach—A methodological comparison between real-time and conventional Delphi studies. *Technol. Forecast. Soc. Chang.* **2011**, *78*, 1681–1694. [CrossRef]
19. Gordon, T.J. Energy forecasts using a "Roundless" approach to running a Delphi study. *Foresight* **2007**, *9*, 27–35. [CrossRef]
20. Aengenheyster, S.; Cuhls, K.; Gerhold, L.; Heiskanen-Schüttler, M.; Huck, J.; Muszynska, M. Real-Time Delphi in practice—A comparative analysis of existing software-based tools. *Technol. Forecast. Soc. Chang.* **2017**, *118*, 15–27. [CrossRef]
21. Adler, M.; Ziglio, E. *Gazing into the Oracle: The Delphi Method and Its Application to Social Policy and Public Health*; Jessica Kingsley Publishers: London, UK, 1996; Volume 5, pp. 3–33.
22. Healy, W.L.; Iorio, R.; Clair, A.J.; Pellegrini, V.D.; Della Valle, C.J.; Berend, K.R. Complications of total hip arthroplasty: Standardized list, definitions, and stratification developed by the hip society. *Clin. Orthop. Relat. Res.* **2016**, *474*, 357–364. [CrossRef]
23. Rampersaud, Y.R.; Sundararajan, K.; Docter, S.; Perruccio, A.V.; Gandhi, R.; Adams, D.; Briggs, N.; Davey, J.R.; Fehlings, M.; Lewis, S.J. Hospital spending and length of stay attributable to perioperative adverse events for inpatient hip, knee, and spine surgery: A retrospective cohort study. *BMC Health Serv. Res.* **2023**, *23*, 1150. [CrossRef]
24. Dindo, D.; Demartines, N.; Clavien, P.-A. Classification of surgical complications: A new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann. Surg.* **2004**, *240*, 205. [CrossRef] [PubMed]
25. Healy, W.L.; Della Valle, C.J.; Iorio, R.; Berend, K.R.; Cushner, F.D.; Dalury, D.F.; Lonner, J.H. Complications of total knee arthroplasty: Standardized list and definitions of the Knee Society. *Clin. Orthop. Relat. Res.* **2013**, *471*, 215–220. [CrossRef] [PubMed]
26. U.S. In Food and Drug Administration (FDA). MAUDE–Manufacturer and User Facility Device Experience. 2019. Available online: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm (accessed on 27th July 2024).
27. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
28. von der Gracht, H.A. Consensus measurement in Delphi studies: Review and implications for future quality assurance. *Technol. Forecast. Soc. Chang.* **2012**, *79*, 1525–1536. [CrossRef]
29. Subong, P.E.; Beldia, M. *Statistics for Research*; Rex Bookstore: Manila, Philippines, 2005.
30. Birko, S.; Dove, E.S.; Özdemir, V. Evaluation of Nine Consensus Indices in Delphi Foresight Research and Their Dependency on Delphi Survey Characteristics: A Simulation Study and Debate on Delphi Design and Interpretation. *PLoS ONE* **2015**, *10*, e0135162. [CrossRef]
31. Raskin, M.S. The Delphi study in field instruction revisited: Expert consensus on issues and research priorities. *J. Soc. Work. Educ.* **1994**, *30*, 75–89. [CrossRef]
32. Rayens, M.K.; Hahn, E.J. Building Consensus Using the Policy Delphi Method. *Policy Politics Nurs. Pract.* **2000**, *1*, 308–315. [CrossRef]
33. Varndell, W.; Fry, M.; Elliott, D. Applying real-time Delphi methods: Development of a pain management survey in emergency nursing. *BMC Nurs.* **2021**, *20*, 149. [CrossRef]
34. Crisp, J.; Pelletier, D.; Duffield, C.; Adams, A.; Nagy, S. The delphi method? *Nurs. Res.* **1997**, *46*, 116–118. [CrossRef]
35. Holey, E.A.; Feeley, J.L.; Dixon, J.; Whittaker, V.J. An exploration of the use of simple statistics to measure consensus and stability in Delphi studies. *BMC Med. Res. Methodol.* **2007**, *7*, 52. [CrossRef]
36. Williams, P.L.; Webb, C. The Delphi technique: A methodological discussion. *J. Adv. Nurs.* **1994**, *19*, 180–186. [CrossRef] [PubMed]
37. Becker, G.E.; Roberts, T. Do we agree? Using a Delphi technique to develop consensus on skills of hand expression. *J. Hum. Lact.* **2009**, *25*, 220–225. [CrossRef] [PubMed]

38. Kalaian, S.A.; Kasim, R.M. Terminating sequential Delphi survey data collection. *Pract. Assess. Res. Eval.* **2012**, *17*, n5.
39. Schmidt, R.C. Managing Delphi Surveys Using Nonparametric Statistical Techniques. *Decis. Sci.* **1997**, *28*, 763–774. [CrossRef]
40. Weir, C.R.; Hicken, B.L.; Rappaport, H.; Nebeker, J.R. Crossing the quality chasm: The role of information technology departments. *Am. J. Med. Qual.* **2006**, *21*, 382–393. [CrossRef]
41. Brender, J.; Ammenwerth, E.; Nykänen, P.; Talmon, J. Factors influencing success and failure of health informatics systems. *Methods Inf. Med.* **2006**, *45*, 125–136.
42. Makovec, U.N.; Goetzinger, C.; Ribaut, J.; Barnestein-Fonseca, P.; Haupenthal, F.; Herdeiro, M.T.; Grant, S.P.; Jácome, C.; Roque, F.; Smits, D.; et al. Developing a medication adherence technologies repository: Proposed structure and protocol for an online real-time Delphi study. *BMJ Open* **2022**, *12*, e059674. [CrossRef]
43. Ahmed, K.; Pendegrass, C.; Aston, W.; Blunn, G. Radiographic Evidence of Bone Changes Around Intraosseous Transcutaneous Amputation Prosthesis: An 11-Year Retrospective Cohort Study. *JPO J. Prosthet. Orthot.* **2024**, *10*, 1097. [CrossRef]
44. Khodyakov, D.; Grant, S.; Denger, B.; Kinnett, K.; Martin, A.; Peay, H.; Coulter, I. Practical Considerations in Using Online Modified-Delphi Approaches to Engage Patients and Other Stakeholders in Clinical Practice Guideline Development. *Patient* **2020**, *13*, 11–21. [CrossRef]
45. Hsu, C.-C.; Sandford, B.A. The Delphi technique: Making sense of consensus. *Pract. Assess. Res. Eval.* **2019**, *12*, 10.
46. Hartman, F.T.; Baldwin, A. Using technology to improve Delphi method. *J. Comput. Civ. Eng.* **1995**, *9*, 244–249. [CrossRef]
47. Diamond, I.R.; Grant, R.C.; Feldman, B.M.; Pencharz, P.B.; Ling, S.C.; Moore, A.M.; Wales, P.W. Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *J. Clin. Epidemiol.* **2014**, *67*, 401–409. [CrossRef] [PubMed]
48. Murphy, M.; Black, N.; Lamping, D.; McKee, C.; Sanderson, C.; Askham, J.; Marteau, T. Consensus development methods, and their use in clinical guideline development. *Health Technol. Assess.* **1998**, *2*, 1–88. [CrossRef]
49. Loughlin, K.G.; Moore, L.F. Using Delphi to achieve congruent objectives and activities in a pediatrics department. *J. Med. Educ.* **1979**, *54*, 101–106. [CrossRef]
50. McKenna, H.P. The Delphi technique: A worthwhile research approach for nursing? *J. Adv. Nurs.* **1994**, *19*, 1221–1225. [CrossRef]
51. Sumsion, T. The Delphi Technique: An Adaptive Research Tool. *Br. J. Occup. Ther.* **1998**, *61*, 153–156. [CrossRef]
52. Green, B.; Jones, M.; Hughes, D.; Williams, A. Applying the Delphi technique in a study of GPs' information requirements. *Health Soc. Care Community* **1999**, *7*, 198–205. [CrossRef]
53. Shang, Z. Use of Delphi in health sciences research: A narrative review. *Medicine* **2023**, *102*, e32829. [CrossRef]
54. Li, Y.; Brånemark, R. Osseointegrated prostheses for rehabilitation following amputation: The pioneering Swedish model. *Der Unfallchirurg* **2017**, *120*, 285. [CrossRef]
55. Juhnke, D.-L.; Beck, J.P.; Aschoff, H.H. Fifteen years of experience with Integral-Leg-Prosthesis: Cohort study of artificial limb attachment system. *J. Rehabil. Res. Dev.* **2015**, *52*, 407. [CrossRef]
56. Al Muderis, M.; Khemka, A.; Lord, S.J.; Van de Meent, H.; Frölke, J.P.M. Safety of osseointegrated implants for transfemoral amputees: A two-center prospective cohort study. *JBJS* **2016**, *98*, 900–909. [CrossRef] [PubMed]
57. Brånemark, R.; Berlin, Ö.; Hagberg, K.; Bergh, P.; Gunterberg, B.; Rydevik, B. A novel osseointegrated percutaneous prosthetic system for the treatment of patients with transfemoral amputation: A prospective study of 51 patients. *Bone Jt. J.* **2014**, *96*, 106–113. [CrossRef] [PubMed]
58. Hagberg, K.; Jahani, S.-A.G.; Kulbacka-Ortiz, K.; Thomsen, P.; Malchau, H.; Reinholdt, C. A 15-year follow-up of transfemoral amputees with bone-anchored transcutaneous prostheses: Mechanical complications and patient-reported outcomes. *Bone Jt. J.* **2020**, *102*, 55–63. [CrossRef] [PubMed]
59. Hagberg, K.; Jahani, S.A.G.; Omar, O.; Thomsen, P. Osseointegrated prostheses for the rehabilitation of patients with transfemoral amputations: A prospective ten-year cohort study of patient-reported outcomes and complications. *J. Orthop. Transl.* **2023**, *38*, 56–64. [CrossRef] [PubMed]
60. Van Eck, C.F.; McGough, R.L. Clinical outcome of osseointegrated prostheses for lower extremity amputations: A systematic review of the literature. *Curr. Orthop. Pract.* **2015**, *26*, 349–357. [CrossRef]
61. Fink, A.; Kosecoff, J.; Chassin, M.; Brook, R.H. Consensus methods: Characteristics and guidelines for use. *Am. J. Public Health* **1984**, *74*, 979–983. [CrossRef]
62. Reetz, D.; Atallah, R.; Mohamed, J.; van de Meent, H.; Frolke, J.P.M.; Leijendekkers, R. Safety and Performance of Bone-Anchored Prostheses in Persons with a Transfemoral Amputation: A 5-Year Follow-up Study. *J. Bone Joint. Surg. Am.* **2020**, *102*, 1329–1335. [CrossRef]